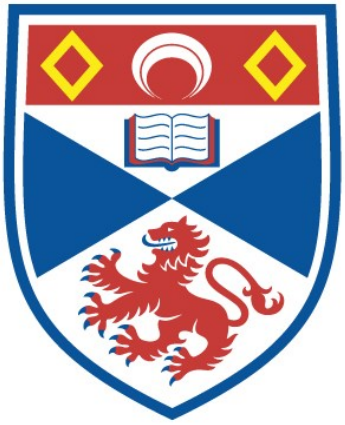
DISTINGUISHING STANDARD FROM MODIFIED GRAVITY IN THE LOCAL GROUP AND BEYOND

Indranil Banik

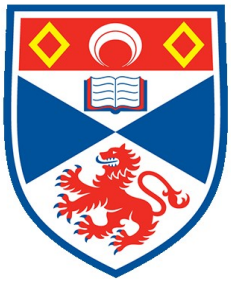
A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2018

Distinguishing Standard from Modified Gravity in the Local Group and Beyond

Indranil Banik



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy (PhD)

at the University of St Andrews

February 2018

Candidate's declaration

I, Indranil Banik, do hereby certify that none of the work contained in the books or papers in this portfolio, which record work substantially carried out by me, has been submitted by me for any degree in any other university.

I was admitted as a research student at the University of St Andrews in September 2014 as a candidate for the degree of PhD.

Date Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this portfolio in application for that degree.

Date Signature of supervisor

In submitting this portfolio to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I certify that I have obtained permission of all copyright holders for this action. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Date Signature of candidate

Distinguishing Standard from Modified Gravity in the Local Group and beyond

Indranil Banik[1]*, Supervisor: Hongsheng Zhao[1]
[1]*Scottish Universities Physics Alliance, University of Saint Andrews, North Haugh, Saint Andrews, Fife, KY16 9SS, United Kingdom*

6 May 2018

ABSTRACT

The works in this portfolio test the hypothesis that it is not possible to extrapolate the Newtonian inverse square law of gravity from Solar System to galaxy scales. In particular, I look into various tests of Modified Newtonian Dynamics (MOND), which posits a modification below a very low acceleration threshold. Although discrepancies with Newtonian dynamics are indeed observed, they can usually be explained by invoking an appropriate distribution of invisible mass known as dark matter (DM). This leads to the standard cosmological paradigm, ΛCDM. I consider how it may be distinguished from MOND using collision velocities of galaxy clusters, which should sometimes be much faster in MOND. I focus on measuring these velocities more accurately and conclude that this test ought to be feasible in the near future.

For the time being, I look at the much nearer and more accurately observed Local Group (LG) of galaxies. Its main constituents – the Milky Way (MW) and Andromeda (M31) – should have undergone a past close flyby in MOND but not in ΛCDM. The fast MW-M31 relative motion around the time of their flyby would have allowed them to gravitationally slingshot any passing LG dwarf galaxies out at high speed. I consider whether there is any evidence for such high-velocity galaxies (HVGs). Several candidates are found in two different ΛCDM models of the LG, one written by a founding figure of the paradigm. The properties of these HVGs are similar to what might be expected in MOND, especially their tendency to lie close to a plane. Being more confident of its validity, I then used MOND to determine the escape velocity curve of the MW over the distance range 8−50 kpc, finding reasonable agreement with the latest observations. I finish by discussing possible future directions for MOND research.

1 INTRODUCTION AND CONTEXT

This PhD is about testing the fundamental physical laws governing the Universe. Currently understood physics is capable of explaining a huge variety of observations. However, the great pillars of modern physics – general relativity and quantum mechanics – are difficult to reconcile (e.g. Carlip 2001). As nature must be internally consistent, it follows that there must be systems whose behaviour can't be accurately predicted using our current incomplete understanding of physics.

Fortunately, there are indeed a few known examples of such systems, some of which were discovered almost a century ago. In particular, the dynamics of galaxies and galaxy clusters require at least one new fundamental assumption (Zwicky 1937; Rubin & Ford 1970). Before considering these observations and why they imply new physics, I will first briefly review our current understanding of the physics which is thought to govern such systems.

The main force at play here is gravity. Although this is intrinsically weaker than the other fundamental forces, it is cumulative in the sense that masses always attract each other. It is the very strength of the electromagnetic force which prevents significant separation of charge and thus ensures this force only acts over short distances in practice. The strong and weak nuclear forces are limited in their range, thus leaving gravity as the dominant influence known to act on large astronomical systems. It is precisely this kind of system which I will focus on, looking in particular at galaxies but also galaxy groups and clusters.

*Email: ib45@st-andrews.ac.uk (Indranil Banik)
hz4@st-andrews.ac.uk (Hongsheng Zhao)

1.1 Gravity in the Solar System

One of the most important advances in our understanding of celestial dynamics came about when Johannes Kepler discovered the eponymous laws of planetary motion in the early 1600s, benefiting from earlier observations by Tycho Brahe. Based on the assumption that planets orbit around the Sun, these empirical laws provided the first indication that Solar System dynamics could be understood with a few simple rules. In particular, Kepler's Third Law – that the orbital period scales as the $\frac{3}{2}$ power of orbital semi-major axis – later proved crucial. Four centuries later, this law is currently at the heart of how data from the Kepler telescope is used to infer orbital properties of exoplanets (Borucki et al. 1997).

Kepler soon realised that this law also applied to the moons of Jupiter, albeit with a different normalisation. Once Newton discovered his laws of motion, consistency with Kepler's Third Law immediately showed that there must be a force towards the Sun with magnitude decreasing as $\frac{1}{r^2}$. This also provided a good explanation for the shapes of cometary orbits (especially Halley's Comet) and for how the Moon remains bound to the Earth. Thus, he soon realised that all massive objects must exert a force towards them in the same way as the Sun, Jupiter and Earth. This gave rise to what we now call Newton's Universal Law of Gravitation. In general, it states that

$$\boldsymbol{g}_N = -\sum_i \frac{GM_i\left(\boldsymbol{r}-\boldsymbol{r}_i\right)}{\left|\boldsymbol{r}-\boldsymbol{r}_i\right|^3} \quad (1)$$

$\boldsymbol{g}_N$ is the predicted acceleration of an object at position $\boldsymbol{r}$ due to the gravity of other masses M_i located at positions $\boldsymbol{r}_i$. The – sign indicates that gravity is attractive. For over two centuries, celestial motions indicated that $\boldsymbol{g}_N = \boldsymbol{g}$ to within observational uncertainties, where $\boldsymbol{g}$ is the actual gravitational field inferred in some less model-dependent way (often based on planetary trajectories). Thus, Equation 1 ruled the heavens until the Industrial Revolution.

Eventually, improved technology enabled more precise observations which highlighted tensions with the theory. These came to a head with the observation of how much light is deflected by the gravitational field of the Sun – the deflection is twice the prediction of Newtonian dynamics (Dyson et al. 1920). Apparently, Newton's laws do not work for relativistic particles like photons. However, there are small deviations even for non-relativistic planets. This is now known to underlie the 'anomalous' precession of Mercury's orbital perihelion by an extra 43" every century. The historical attempt to explain this small but statistically significant discrepancy is excellently reviewed in Ruskin (2017), highlighting several analogies with the ongoing missing mass vs modified gravity debate at the heart of this thesis.

In addition to observational discrepancies, it was also necessary to reconcile Newtonian gravity with the Special Theory of Relativity (Einstein 1905). The latter precluded instantaneous action at a distance, even though this is how the former works. These issues were resolved using the General Theory of Relativity (Einstein 1915). With this in hand, it seemed that all Solar System motions could be adequately explained.

General Relativity often yields very similar predictions to Newtonian gravity. This is especially true in galaxies, which are the focus of this article. Thus, I will treat the predictions of the latter as equal to those of the former (except when discussing gravitational lensing). This is because Newtonian gravity is much simpler to handle than General Relativity. Although the latter is important for cosmological-scale problems, I do not directly address such large scales. In particular, I only consider systems whose mean density is much larger than that of the Universe as a whole, making the system effectively decoupled from the large-scale cosmic expansion (Hubble flow, Hubble 1929).

1.2 Gravity beyond the Solar System – the discovery of acceleration discrepancies

In the past century, observations beyond the Solar System became increasingly accurate. These showed a remarkable phenomenon that I shall call 'acceleration discrepancies'. Although it is currently not possible to directly observe accelerations in systems much larger than the Solar System, one can reasonably assume that a star in a rotating disk galaxy has a centripetal acceleration of $\frac{v^2}{r}$. This requires careful observation of the galaxy to be sure that it really is a rotating disk. Fortunately, this can be confirmed with only minimal assumptions based on the line of sight ('radial') velocity of its different parts. These motions cause a Doppler shift in the wavelengths of spectral lines that are nowadays measurable using integral field unit spectroscopy. In the case of the Large Magellanic Cloud, the rotation of the galaxy can be seen directly using proper motions (van der Marel & Kallivayalil 2014). In this way, it is possible to obtain an observational estimate of the acceleration that makes few assumptions, especially with regards to the gravity theory.

As we do have such theories, this opens the possibility of testing them much more thoroughly using the latest Galactic and extragalactic observations. To be useful, theories of gravitation need to predict the acceleration based on the mass distribution (e.g. using a procedure similar to Equation 1). With some assumptions, we can convert observed light into an idea of how the mass is distributed in a particular system, thus determining the expected $\boldsymbol{g}\left(\boldsymbol{r}\right)$.

When Newtonian gravity is used to do this, the predicted acceleration often falls far short of the observed value. An early example of this was in the Coma Cluster of galaxies, where Fritz Zwicky found the need for $\sim 100\times$ more matter than suggested by its observed brightness (Zwicky 1937). It was eventually realised that much of this mass exists as hot gas, which is in fact very bright – but in X-rays, inaccessible to observations by ground-based telescopes (Sarazin 1986).

Another major acceleration discrepancy was found in the rotation curves of disk galaxies. Beyond the majority of their visible matter, Equation 1 implies the rotation speed should decrease as $v_c \propto \frac{1}{\sqrt{r}}$, the rotation curve version of Kepler's Third Law. However, observed rotation curves tend to remain flat out to large distances. An early indication of this came from the nearest large external galaxy, Andromeda (M31, Babcock 1939). Later work confirmed that M31 indeed rotates much faster in its outer parts than can be expected on the basis of its visible mass (Rubin & Ford 1970). This was later confirmed with radio observations of the 21 cm hyperfine transition of neutral hydrogen (Roberts & Whitehurst 1975). Such observations also indicated flat rotation curves for several other galaxies (Rogstad & Shostak

1972). Radio observations were important because they extend out to much larger radii than optical measurements. This is due to star formation requiring a threshold gas density. At sufficiently large distances, the exponentially declining gas density (Freeman 1970) falls below this threshold, leaving the outer parts of galaxies with very few stars.

If rotation curves are flat, the discrepancy with a Keplerian decline should become more pronounced at a larger distance. This should be easily detectable when considering the forces *between* galaxies rather than the internal forces within them. In this regard, an important constraint is provided by the dynamics of the Milky Way (MW) and M31, galaxies which are $\sim$0.8 Mpc apart (McConnachie 2012). The basic idea is that they must have started receding from each other shortly after the Big Bang. However, they are presently approaching each other at $\sim$110 km/s, as inferred from the observed radial velocity of M31 (Slipher 1913) corrected for the motion of the Sun within the MW (Schmidt 1958). Therefore, the gravitational attraction between the galaxies must have been strong enough to turn their initial recession around (Kahn & Woltjer 1959).[1] Using this constraint (known as the timing argument), it was found that the total mass in the MW & M31 needed to be $\sim 4\times$ the observed matter in them.

Although this result was surprising, there were other reasons to suppose that the MW & M31 are more massive than might be expected from imaging of their visible disks. With the advent of computers, Hohl (1971) used N-body simulations to show that self-gravitating disks are unstable, rapidly becoming dynamically hot (i.e. developing non-circular motions comparable to the circular rotation speed). The instability develops over only a few orbital periods, whereas the Universe is $\sim 40\times$ older than the orbital period of the Sun (McMillan 2017). Thus, observed spiral galaxies can't be self-gravitating and must be surrounded by a dynamically hot halo. Moreover, this halo has to dominate the mass of the galaxy. As no such component is seen, Ostriker & Peebles (1973) suggested that it is dark.

1.3 The Massive Compact Halo Object hypothesis

One possibility for this dark matter (DM) was a large number of as yet undetected very faint stars or stellar remnants around each galaxy (Carr 1994). This theory of massive compact halo objects (MACHOs) could be tested using gravitational microlensing searches (Kerins & Carr 1995). The basic idea is that a massive object would occasionally appear to pass very close to a star on our sky (Refsdal 1966). This alignment would cause the foreground mass to gravitationally deflect light from the background star, which would therefore appear to brighten and then fade.[2]

The obvious problem with searching for such microlensing events is that the true luminosity of a star can change. Generally, this would be associated with a change in its temperature. This would alter the colour of the star $\equiv$ the ratio of its fluxes in two different wavebands. However, gravitational microlensing equally affects photons of all wavelengths. This allows us to distinguish between microlensing and intrinsic variability by observing in two or more wavelength bands (Paczynski 1986).

Using these ideas, the EROS collaboration conducted a careful search for microlensing events. This involved continuous monitoring of 7 million stars in two fields of view towards the Large and Small Magellanic Clouds over a period of 6.7 years. Instead of the $\sim$39 events expected under the MACHO hypothesis, only 1 candidate event was found (Tisserand et al. 2007). Similar results had already been reached several years earlier (Alcock et al. 2000). As a result, it has become clear that MACHOs almost certainly do not have enough mass to account for the acceleration discrepancies in our Galaxy or to stabilise its disk if it obeys Newtonian dynamics.

1.4 Non-baryonic dark matter

This leads to several possibilities, none of which are based solidly on existing laws of physics. The most popular idea is to maintain the assumption of a large amount of mass in the outskirts of galaxies. This DM hypothesis is one of the key pillars of the currently prevailing cosmological paradigm (ΛCDM, Ostriker & Steinhardt 1995). Cold gas in the amounts required would easily be detected and would in any case likely have clumped into MACHOs, contradicting microlensing observations (Section 1.3). Although a small amount of hot gas is expected and has indeed been detected around the MW (Nicastro et al. 2016), this can't constitute all of the DM. Thus, one needs to assume that the DM is not composed of baryons at all.

This leads to the present situation where no known fundamental particle has the properties required of the DM. Thus, it is thought to consist of an undiscovered stable particle, or at least one with a decay time longer than the age of the Universe (e.g. Steigman & Turner 1985, and references therein). The leading contender is a weakly interacting massive particle (WIMP, Griest 1993), though a much lower mass axion could also work (Kamionkowski 1998).

Multi-decade searches for a WIMP have now ruled out a substantial part of the parameter space that was thought to be feasible before the searches started (e.g. Fermi-LAT Collaboration 2015; LUX Collaboration 2017; PandaX-II Collaboration 2016). Moreover, an important motivation for the WIMP hypothesis is that nature might respect a new fundamental symmetry called supersymmetry (Jungman et al. 1996). This predicts a plethora of new particles. However, recent null results from the Large Hadron Collider argue against the simplest forms of supersymmetry (ATLAS Collaboration 2015).

Less attention has been paid to the possibility of axion DM, though this has started to change recently due to null detections of WIMPs (Baer et al. 2015). Axions may be easier to search for as they interact with a strong magnetic field (Sikivie 1983). As neutron stars indeed have very strong magnetic fields, this has allowed some constraints to be placed on axion properties if they are ubiquitous enough to comprise the DM (Berenji et al. 2016). Although a promising start, this leaves open most of the axion mass range

[1] Deviations of velocity $\boldsymbol{v}$ from a pure Hubble expansion ($\boldsymbol{v} \equiv H\boldsymbol{r}$) are called peculiar velocities ($\boldsymbol{v}_{pec} \equiv \boldsymbol{v} - H\boldsymbol{r}$), where $\boldsymbol{r}$ is used for position and $H \equiv \frac{\dot{a}}{a}$ is the logarithmic time derivative of the cosmic scale-factor $a\,(t)$.

[2] There would be two apparent images of the star, but in microlensing these are – by definition – unresolved. If they are resolved, then the lensing is said to be 'strong'.

calculated by Borsanyi et al. (2016). In fact, this range is difficult to probe by the Axion Dark Matter Experiment (Duffy et al. 2006), one of the longest-running searches for axion DM. Thus, the acceleration discrepancies may yet be resolved using axions. Until then, it is prudent to consider other possibilities.

1.5 Modified Newtonian Dynamics

Just as we do not yet have a complete understanding of particle physics, so also we do not yet understand gravity. Therefore, another possibility is to suppose that the acceleration discrepancies are caused by a breakdown of Newtonian gravity in the relevant systems. This rather old idea was alluded to by Zwicky (1937) in the same paper that first reported significant acceleration discrepancies. There, Zwicky suggested that the inverse square law of gravity might break down at large distances.

As more observational data was gathered, certain patterns in the acceleration discrepancy became apparent – where it appeared and where it did not. In this respect, a crucial discovery was the Tully-Fisher Relation concerning the dynamics of spiral galaxies (Tully & Fisher 1977). Eventually, Mordehai Milgrom realised that the important physical parameter is not the size of a system but the typical acceleration within it (Milgrom 1983). If modified gravity is the answer, then Newtonian gravity needs to break down below an acceleration scale a_0. This theory of Modified Newtonian Dynamics (MOND) assumes that the gravitational field strength g at distance r from an isolated point mass M transitions from the usual inverse square law (Equation 1) at short range to

$$g = \frac{\sqrt{GMa_0}}{r} \quad \text{for} \quad r \gg \sqrt{\frac{GM}{a_0}} \tag{2}$$

a_0 is a fundamental acceleration scale of nature which must have an empirical value close to 1.2×10^{-10} m/s^2 to match galaxy rotation curves (McGaugh 2011).

In the 1990s, another unexplained acceleration was observed – that of the whole Universe. Instead of slowing down due to the attractive effect of gravity, the cosmic scale-factor $a(t)$ seemed to be speeding up ($\ddot{a} > 0$, Riess et al. 1998). This could be fit into the context of General Relativity by reintroducing the cosmological constant term Λ, a direct coupling between the metric and Ricci curvature tensors. This 'dark energy' can be viewed as a uniform energy density fundamental to the fabric of spacetime itself. Considering the behaviour of quantum systems, this makes some sense – such systems have a zero point energy due to inherent uncertainty in field strengths and their time derivatives (e.g. in the position and velocity of a particle). Thus, a pendulum can never be exactly at the bottom and have zero velocity. Consequently, the energy of the pendulum must be slightly above the classical minimum.

Similarly, an apparently empty region of spacetime must have some value for quantities such as the electric field strength. Although it might be 0 classically, this is no longer feasible quantum mechanically – there must be some uncertainty. For this reason, it is possible that spacetime itself has a minimum (zero-point) energy density associated with it – a cosmic ground state.

This quantum-mechanical phenomenon seems to be having a significant effect on the expansion rate history of the Universe. This raises the question of whether there are other circumstances in which quantum effects might force us to revise our classical (non-quantum) expectations for the motions of astrophysical objects. A possibly useful analogy could be drawn with a gas – at high enough temperatures, it behaves classically. However, quantum effects become important at low temperatures, when large-scale properties such as the heat capacity start to behave differently. Instead of being temperature-independent, this decreases with temperature due to the 'freezing out' of quantised degrees of freedom that ultimately underlie heat capacity. A rough estimate of when this occurs (i.e. when equipartition of energy breaks down) can be found by equating the classical result for the mean energy of each particle with the Fermi energy.

The details of how quantum mechanics works with gravity are still unclear. Classically, the energy density in a gravitational field is given by

$$u = -\frac{g^2}{8\pi G} \tag{3}$$

In a remarkable coincidence called the cosmic coincidence of MOND, a_0 is comparable to the value of g at which this equation yields an energy density equal in magnitude to the dark energy density $u_\Lambda = \rho_\Lambda c^2$ implied by the accelerating expansion of the Universe (Riess et al. 1998). Thus,

$$\frac{g^2}{8\pi \mathrm{G}} < u_\Lambda \quad \Leftrightarrow \quad g \lesssim 2\pi \mathrm{a}_0 \tag{4}$$

This strongly suggests that MOND is simply an empirical way of capturing deviations from classical gravity which arise due to quantum effects (Milgrom 1999). After all, u_Λ is likely a quantum mechanical effect – if it dominates the energy density in a particular region of spacetime, then quantum gravity effects could well be important. Assuming that our classical gravity theories are only approximations to the true quantum gravity theory underlying nature, it would not be surprising if our existing theories failed at accelerations $\lesssim a_0$ but worked at higher accelerations. Indeed, there are some specific suggestions for how quantum gravity might work which yield MOND-like behaviour at low accelerations (e.g. Pazy 2013; Verlinde 2017; Smolin 2017).

MOND was originally formulated as a non-relativistic theory, only to be applied in systems where Newtonian gravity and General Relativity meant much the same thing. This covers internal motions of galaxies and should cover forces between nearby galaxies. To describe these situations, a modified version of the usual Poisson Equation of Newtonian gravity is used (Bekenstein & Milgrom 1984).

$$\nabla \cdot [\mu(g)\boldsymbol{g}] = -4\pi G\rho \quad \text{where} \tag{5}$$

$$g \equiv |\boldsymbol{g}| \quad \text{and} \tag{6}$$

$$\mu(g) = \frac{g}{g + a_0} \tag{7}$$

Here, I used the simple interpolating function $\mu(g)$ to capture how nature transitions between the Newtonian ($g \gg a_0$) and deep-MOND ($g \ll a_0$) regimes (Famaey & Binney 2005). This fairly gradual transition works very well with high-precision kinematic data from our own Galaxy (Iocco et al. 2015) and from a large sample of $\sim$6000 elliptical galaxies probing accelerations up to $\sim 30a_0$ (Chae et al.

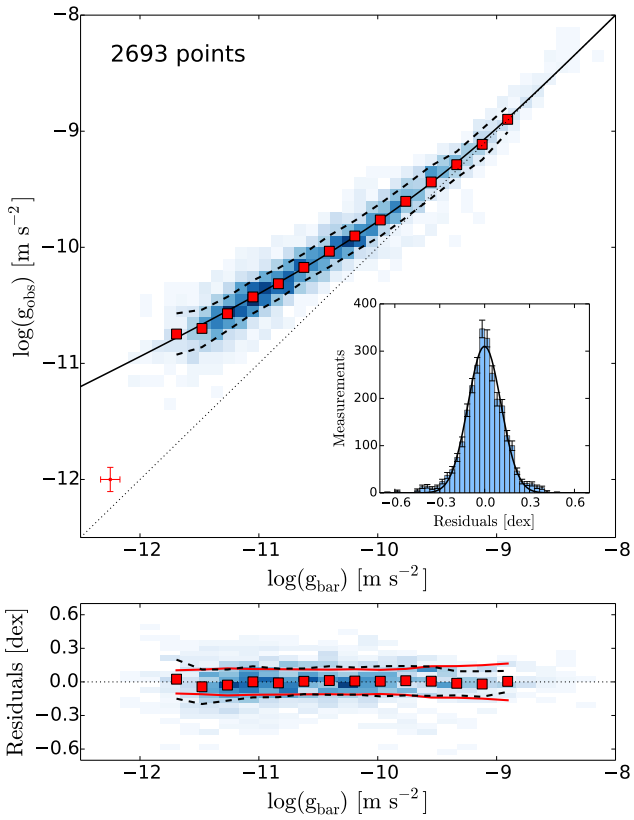


Figure 1. The relation between the actual acceleration g_{obs} in rotationally supported disk galaxies and the prediction g_{bar} of Newtonian gravity using their visible matter. Data from 153 galaxies (yielding 2,693 data points) are shown here as a 2D histogram, with darker shades of blue used to represent more common bins. The typical error budget is shown in red towards the bottom left. Figure from McGaugh et al. (2016).

2017, figure 2). In spherical symmetry, it implies that the true gravity g can be obtained from the Newtonian gravity g_N using

$$g = \frac{g_N}{2} + \sqrt{\left(\frac{g_N}{2}\right)^2 + g_N a_0} \quad (8)$$

MOND has been remarkably successful at predicting rotation curves of disk galaxies merely by applying Equation 5 to their distribution of visible mass (Famaey & McGaugh 2012). This works because there is a very tight correlation between the actual accelerations in such galaxies and the predictions of Newtonian gravity over $\sim$5 orders of magnitude (dex) in mass ($10^7 M_\odot - 10^{12} M_\odot$) and $\sim$2 dex in surface brightness (Figure 1). This radial acceleration relation (RAR) is underpinned by mass estimates based on near-infrared photometry collected with the Spitzer Space Telescope (Lelli et al. 2016), taking advantage of reduced variability in stellar mass-to-light (M/L) ratios at these wavelengths (Bell & de Jong 2001; Norris et al. 2016). The kinematics are estimated using only the most reliable rotation curves (Lelli et al. 2016, section 3.2.2). The tightness of the RAR in the face of observational uncertainties is perhaps the clearest indication yet that our current understanding of gravity breaks down at very low acceleration.

In this context, it would be easy to explain why Newtonian gravity is off by the same factor close to a low mass galaxy and far from a more massive galaxy, as long as g_N is equal at both positions. However, in a DM context, this requires a tight correlation between each galaxy's rotation curve shape, DM halo scale radius and mass such that $\lesssim 10^{-5}$ of the available phase space volume is actually filled (Salucci et al. 2007). Those authors noted that "theories of the formation of spirals do not trivially imply the existence of such a surface that underlies the occurrence of a strong dark-luminous coupling." More recent investigations continue to have difficulty explaining such correlations using collisionless DM, the standard version of ΛCDM (Salucci & Turini 2017).

Such correlations are intrinsic to MOND, which therefore predicts a global scaling relation between the asymptotic rotation curve v_f of a galaxy and its mass M. Beyond the bulk of the visible light from any galaxy, it can be well-approximated as a point mass in the deep-MOND regime ($g \ll a_0$), thus making Equation 2 valid. As the centripetal acceleration is $\frac{v^2}{r}$ and g is also $\propto \frac{1}{r}$, the rotation curve of every galaxy should eventually go flat at a level

$$v_f = \sqrt{rg} \quad (9)$$

$$= \sqrt[4]{GMa_0} \quad (10)$$

Observed rotation curves are indeed asymptotically flat, with v_f related to the baryonic mass of a galaxy according to the empirical Baryonic Tully-Fisher Relation (BTFR). Figure 3 of Famaey & McGaugh (2012) − reproduced in my Figure 2 − shows that it agrees very well with Equation 10, the MOND prediction for the BTFR. In the context of ΛCDM, this agreement is puzzling. It implies a baryon:DM ratio that varies between galaxies and is often much lower than in the Universe as a whole (∼5, Planck Collaboration XIII 2016). This is only possible if a substantial fraction of the baryons have been lost from most spiral galaxies − loss of the sub-dominant baryons would not much affect the total mass and rotation curve, but it would reduce the baryon:DM ratio. Recent work on heavy element abundances in galaxies strongly argues against such large amounts of baryons being lost (Peng et al. 2015).

Supposing nonetheless that this is possible, the DM would evolve quite differently to the baryons as DM can't radiate. Thus, while supernovae (SNe) can heat up gas and eject it from a galaxy, the DM could not be directly heated by radiation from SNe in this way. As events like SNe are to some extent stochastic, one expects different relative amounts of baryons and DM in different systems.

As an example, I consider the gas fractions of galaxies with equal baryonic mass. Gas-rich and gas-poor galaxies must have evolved differently. The gas-poor galaxy most likely had much more SNe and ejected a larger fraction of its baryons. However, as the DM component dominates at large radii, v_f must be a property of its distribution. Therefore, one expects the galaxy with the lower gas fraction to have started out with more baryons i.e. in a more massive DM halo. This would imply a higher v_f than for the gas-rich galaxy. Yet, there is no correlation between the gas fraction and how far off a galaxy is from satisfying Equation 10, i.e. the BTFR residual (Lelli et al. 2017b, figure 4). This figure also shows that the BTFR residual is uncorrelated

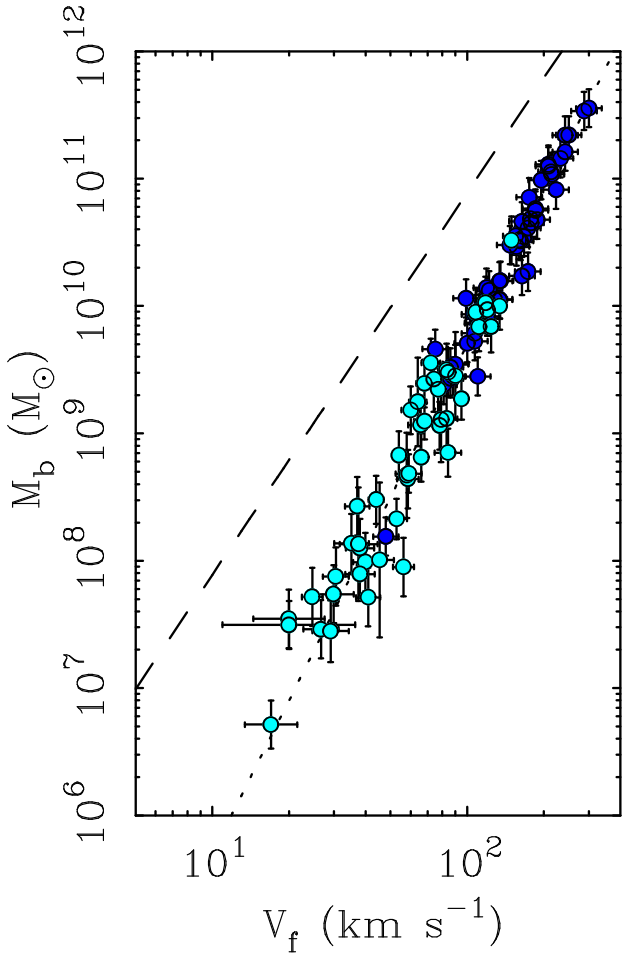


Figure 2. Asymptotic rotation velocity v_f as a function of the total (stellar plus gas) baryonic mass M_b for $\sim$80 galaxies. The dashed line shows the expected trend using Newtonian gravity and a constant ratio of 5:1 between dark matter and baryons, the ratio required by ΛCDM for the Universe as a whole (Planck Collaboration XIII 2016). The dotted line shows the prediction of MOND (Equation 10). Dark blue points are star-dominated galaxies while light-blue points are gas-dominated ones. Figure from Famaey & McGaugh (2012).

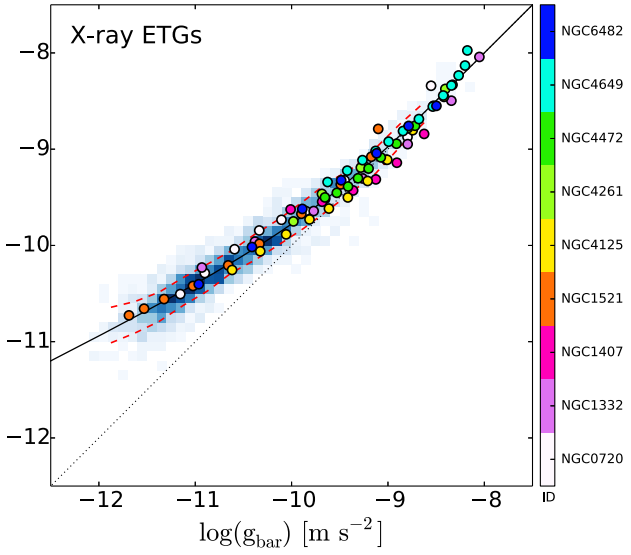


Figure 3. Similar to Figure 1 but showing results for elliptical galaxies as coloured dots, with different colours used to represent different galaxies. The data is based on Chandra and XMM-Newton X-ray observations. The relation for spiral galaxies (Figure 1) is shown in the background using various shades of blue (more frequent regions shown in a darker shade). Figure from Lelli et al. (2017b), which shows that similar results are obtained using the 21 cm neutral hydrogen line (see their figure 8).

with galaxy mass, size and surface density (only the last is a direct consequence of how MOND works as surface densities are related to accelerations, but total mass by itself need not be).

The environment of each galaxy should also play some role. But even in the same conditions, it is inevitable that star formation and SNe feedback is somewhat stochastic, especially in dwarf galaxies. This makes it all the more surprising that there is no evidence for any deviation from Equation 10 over $\sim$5 orders of magnitude in baryonic mass and a similar range in surface density. In fact, observations constrain any possible intrinsic scatter to $\lesssim$ 0.05 dex (12%). Nearly 40 years after Equation 10 was first proposed (Milgrom 1983), it has remained consistent with rotation curve observations.

A relation roughly like the RAR should arise in ΛCDM because lower mass DM halos have shallower gravitational potential wells. This should make it easier for baryons to be ejected via energetic processes like SNe feedback. Still, the tightness of the observed RAR is difficult to explain in this way (Desmond 2017a,b). Some attempts have been made to do so (e.g. Keller & Wadsley 2017), but so far these have investigated only a very small range of galaxy masses and types. In these limited circumstances, there does seem to be a tight correlation of the sort observed. However, a closer look reveals that several other aspects of the simulations are inconsistent with observations (Milgrom 2016). For example, the rotation curve amplitudes are significantly overestimated in the central regions (Keller et al. 2016, figure 4). This issue was recently revisited by Tenneti et al. (2018), who found that it was possible to get a tight RAR but with the wrong low-acceleration behaviour ($g \propto {g_N}^{0.7}$ rather than the observed ${g_N}^{0.5}$) and too high a transition acceleration above which $g \rightarrow g_N$.

Although the disk galaxy RAR may eventually be accounted for in ΛCDM, it has recently become clear that elliptical galaxies follow the same RAR as spirals (Figure 3). This poses additional problems for ΛCDM because feedback would almost certainly work quite differently in spiral and elliptical galaxies, surely leaving them with different proportions of dark and visible matter.

Galactic gravitational fields can also be probed based on how they deflect light from background galaxies. The lensing is said to be strong if it leads to two distinct images. MOND and General Relativity are expected to have the same relation between gravitational field and light deflection (Chiu et al. 2006), at least in the tensor-vector-scalar (TeVeS) relativistic extension of MOND (Bekenstein 2004). With this assumption, Tian & Ko (2017) showed that MOND works

rather well with the Sloan Lensing Advanced Camera for Surveys strong lens sample of elliptical galaxies (Bolton et al. 2008).

If the background galaxy is not multiply imaged but only appears mildly distorted, then this is termed weak gravitational lensing. It has recently become possible to detect weak lensing by galaxies using stacking techniques (Brimioulle et al. 2013). As this requires a large number of background galaxies to detect, it can't probe too close to the lensing galaxy (within $\sim$50 kpc). However, it can probe much further out than rotation curves, reaching a typical distance of 250 kpc. Thus, weak lensing and rotation curves are complementary probes of galactic gravitational fields.

50 kpc is much larger than the typical extent of galaxies and their MOND radii (e.g. the MOND radius of the MW is 9 kpc while its disk scale length is 2.15 kpc, Bovy & Rix 2013). Thus, the MOND-predicted gravitational field in the relevant region can be well approximated by Equation 2. This leads to the simple prediction that a mass M deflects any sufficiently distant ray of light through an angle

$$\Delta\phi \;=\; \frac{2\pi\sqrt{GMa_0}}{c^2} \;=\; 2\pi\left(\frac{v_f}{c}\right)^2 \qquad (11)$$

Milgrom (2013) showed that this rather simple expectation is consistent with the weak lensing data collected by Brimioulle et al. (2013). Importantly, Equation 11 works around both spiral and elliptical galaxies if we assume that foreground galaxies with redder colours are typically ellipticals which have a higher M/L than the bluer spirals.

Individual rotation curves often reveal additional insights. This is especially true for low surface brightness galaxies, which exhibit larger acceleration discrepancies. For example, NGC 1560 has large discrepancies at all radii (Broeils 1992). It has a pronounced dip in its baryonic radial density profile at $\sim$5 kpc. In ΛCDM, this would not affect the overall density of matter very much, because the vast majority of the matter must be dark. As the DM can't radiate and cool, it would not have settled into a disk (unlike the baryons). In any case, the DM needs to remain in a spheroidal halo to explain the stability of the observed disk (Ostriker & Peebles 1973).

Being dynamically hot, the DM would have a smooth distribution that is hardly affected by a sharp dip in the surface density of the sub-dominant baryons. As this smooth DM component needs to dominate the mass of the galaxy, one might expect the resulting gravitational field to be smooth, yielding a smooth rotation curve. However, there is a sharp dip in the rotation curve corresponding to that in the distribution of baryons (Gentile et al. 2010).

NGC 1560 is just one example of features in the rotation curve corresponding to features in the underlying distribution of baryons. In fact, the correspondence is almost one-to-one (Sancisi 2004). While unsurprising when the baryons dominate the gravitational field, this is surprising when the smooth DM component is supposedly dominant.

There are other hints that the gravitational field in low surface brightness galaxies is actually dominated by their baryonic disk. Such galaxies often show spiral structure (e.g. McGaugh et al. 1995). This is probably the result of a self-gravitating instability in the disk (Lin & Shu 1964). However, the disk can't be self-gravitating if the mass of the galaxy is dominated by DM, even in the inner parts. This suggests that the large acceleration discrepancies in such systems are caused by matter distributed within the baryonic disk.

In MOND, the stability problem of disk galaxies is resolved by modifying the gravitational field equation to Equation 5 (Brada & Milgrom 1999). Roughly speaking, this is because $\boldsymbol{g}$ is sub-linearly affected by the underlying matter density (in Equation 2, $g \propto \sqrt{M}$ rather than the Newtonian scaling $g \propto M$). Thus, an enhancement to the density has a smaller effect on $\boldsymbol{g}$, thereby limiting the tendency of material to be attracted to the overdensity and enhance it further. This makes it more likely that other mechanisms (e.g. pressure) will stabilise the perturbation.

In this way, MOND might be able to confer on disk galaxies just the right amount of stability – neither too much to 'smother' spiral structure nor too little to easily let the galaxy evolve into a pressure-supported system. The stabilising mechanism would not work for disk galaxies with a sufficiently high surface density as these would be in the Newtonian regime (if the central surface density $\pi G\Sigma_0 \gg a_0$). This may explain why such disks do not exist in nature (Freeman 1970; McGaugh 1996). If these disks formed at all, perhaps their central parts were unstable and formed into a bulge within a few dynamical times.

In ΛCDM, the large sizes of the DM halos around individual galaxies would lead to frequent collisions between them (White & Frenk 1991). These collisions would cause mergers through the process of dynamical friction, the tendency of a massive object moving through a field of lower mass particles to gravitationally focus those particles behind it (Chandrasekhar 1943). This wake then exerts a gravitational force on the object which created it, slowing it down.

As a result, soon after two similarly massive DM halos merge, their central galaxies should also merge (Privon et al. 2013). The timescale for this would typically be the crossing time of the DM halo, much less than the age of the Universe. As a result, spiral galaxies ought to be quite rare. Even spiral galaxies that avoid major mergers should still have significant bulges due to star clusters losing angular momentum to the surrounding DM halo via dynamical friction (Noguchi 1999; Combes 2014). In theory, strong stellar feedback could disrupt the cluster quickly, preventing such an inspiral. However, recent observations suggest that star clusters can survive for hundreds of Myr (Zanella et al. 2015). This leads to a major contradiction with recent observations indicating a high fraction of bulgeless disk galaxies (Kormendy et al. 2010). To highlight the seriousness of the problem, Martig et al. (2012) stated that "no simulation has ever been able to produce a MW-mass bulgeless galaxy" in a ΛCDM context.

The high frequency of not just disk galaxies but bulgeless disks poses problems to the idea that galaxies are surrounded by massive DM halos capable of exerting dynamical friction. Such friction arises regardless of the exact nature of the DM particles, as long as they exert gravity (as they must, if halos of such particles cause the acceleration discrepancies). Without such halos, dynamical friction would be reduced and galaxies would be much smaller, making them less likely to collide. This is almost certainly why bulges are much less prevalent and much smaller in MOND than in ΛCDM (Combes 2014).

1.6 Satellite planes and tidal dwarf galaxies

ΛCDM faces another problem with the detailed properties of galaxies in the Local Group (LG). Wide field surveys such as the Sloan Digital Sky Survey (York et al. 2000) and the Pan-Andromeda Archaeological Survey (McConnachie et al. 2009) have shown that the satellite galaxies of the MW are preferentially located in a thin co-rotating planar structure (Pawlowski & Kroupa 2013). The same is also true of Andromeda (Ibata et al. 2013), though co-rotation can't be definitively confirmed without proper motions.

It appears very unlikely that these structures formed quiescently (Pawlowski et al. 2014; Ibata et al. 2014b). For the MW, filamentary infall is considered unlikely because this would imply its satellites had very eccentric orbits, contrary to observations (Angus et al. 2011). These require the accretion to have been long ago in order to give enough time to circularise the orbits via dynamical friction against the Galactic DM halo. However, interactions between satellites and numerous DM halos that are thought to surround the MW would cause the dispersal of any initially thin disk of satellites (Klimentowski et al. 2010). A similar phenomenon would be expected around M31 (Fernando et al. 2018). Even if the number of subhalos was smaller than predicted by ΛCDM, the triaxial nature of the potential would lead to any disk-like structure spreading out on a timescale of ~ 5 Gyr unless it was fortuitously aligned with a symmetry axis of the potential (Bowden et al. 2013; Fernando et al. 2017). This issue is less serious in MOND as the matter distribution is much more concentrated, leading to a nearly spherical potential beyond ~ 40 kpc (Figure 26).

After careful consideration of several proposed explanations for why primordial satellites now lie in a thin plane, Pawlowski et al. (2014) concluded that none of them agreed with observations for either the MW or M31. It was later shown that baryonic effects are unlikely to provide the necessary anisotropy if one sticks to a primordial origin for the satellites (Pawlowski et al. 2015). This issue was revisited by performing a high-resolution ΛCDM hydrodynamical simulation of a MW analogue in a cosmological context (Maji et al. 2017). Although this unpublished article claimed that the results were consistent with observations, it has recently been shown that this is not the case (Pawlowski et al. 2017). Those authors showed that the satellite galaxy distribution of Maji et al. (2017) was consistent with isotropy. However, the actual MW satellite system is inconsistent with isotropy at more than 5σ once the survey footprint is taken into account (Pawlowski 2016).

More recent hydrodynamical ΛCDM simulations also fail to yield highly flattened satellite systems like those observed around the MW and M31 (Ahmed et al. 2017). The mild flattening in these simulations might not even be related to baryonic effects as similar results arise in DM-only simulations (Garaldi et al. 2018). In any case, it is difficult to see how baryonic effects like radiative cooling can explain a ~ 200 kpc-wide plane of primordial satellites composed mostly of DM.

This raises the possibility that most LG satellites are not primordial − perhaps they formed as second-generation tidal dwarf galaxies (TDGs) during an ancient galactic interaction (Kroupa et al. 2005). After all, we do see galaxies forming from material pulled out of interacting progenitor galaxies (e.g. in the Antennae, Mirabel et al. 1992). This naturally leads to anisotropy because the tidal debris tend to be confined within the common orbital plane of the interacting progenitor galaxies.

TDGs form by self-gravitating collapse, requiring a high density. This is easy to obtain by tidally perturbing baryons originally on near-circular orbits in a rotating disk, leading to a thin dense tidal tail. However, the DM halos hypothesised to surround galaxies need to be dynamically hot (Hohl 1971). Tidally perturbing this rather sparse (albeit massive) halo would yield only a very low density, insufficient to reach the threshold for Jeans instability. Consequently, TDGs should be free of DM (Barnes & Hernquist 1992; Wetzstein et al. 2007). Their rather low escape velocity also precludes them from subsequently accreting significant amounts of DM out of their host galaxy's halo.

Thus, a surprising aspect of LG satellites is their high Newtonian dynamical masses compared to their low luminosities (e.g. McGaugh & Milgrom 2013). These M/L ratios are calculated assuming dynamical equilibrium, an assumption which could be invalidated by tides from the host galaxy. However, tides are likely not strong enough to do this (McGaugh & Wolf 2010, figure 6). As DM is unlikely to be present in these systems, some other explanation must be found for their high inferred M/L ratios.

This is true even in unconventional models of DM where it has significant non-gravitational interactions with baryons (Famaey et al. 2018). This model is designed to explain the RAR and arguably can do so in the visible regions of both spiral and elliptical galaxies. However, their table 1 shows that TDGs are expected to be free of DM and thus follow standard Newtonian behaviour. Similarly, the MW and M31 satellite planes are rather extended (e.g. Pawlowski & Kroupa 2013, figure 2) and would very likely reach beyond the hypothetical superfluid DM halos of their host galaxies (Berezhiani & Khoury 2016; Khoury 2016). This would cause the more distant MW and M31 satellite plane members to follow Newtonian dynamics.

Without DM, the strong self-gravity needed to maintain high internal velocity dispersions arises most naturally from a modification to gravity. In the context of the most widely investigated such model, the MW and M31 would have undergone an ancient close flyby ~ 8 Gyr ago (Banik & Zhao 2018b, figure 4). Initial N-body simulations of this flyby in MOND suggest that this is a plausible scenario, though it is not yet clear if it can match LG properties in detail (Bílek et al. 2017).

A past encounter with M31 might naturally account for the MW thick disk (Gilmore & Reid 1983), a structure which seems to have formed fairly rapidly from its thin disk (Hayden et al. 2015) 9 ± 1 Gyr ago (Quillen & Garnett 2001). More recent investigations suggest that there was a burst of star formation at that time (Snaith et al. 2014, figure 2). The star formation rate of M31 also appears to rise sharply for lookback times $\gtrsim 9$ Gyr (Williams et al. 2017b, figure 12). The disk heating which likely formed the MW thick disk appears to have been stronger in its outer parts, characteristic of a tidal effect (Banik 2014). This may explain why the Galactic thick disk has a longer scale length than its thin disk (Jurić et al. 2008; Jayaraman et al. 2013).

One possible objection to this theory is that the heavy element abundances of the planar M31 satellites seem rather

similar to those outside its satellite plane (Collins et al. 2015). One might expect there to be a difference if some M31 satellites formed from material already enriched by virtue of being within the disk of a massive galaxy (M31) while others formed primordially. However, this difference becomes very small if the MW-M31 interaction was a very long time ago. This is because there would have been little time to enrich the gas in the M31 disk. M31 would very likely have been much more gas-rich than at present, diluting any heavy elements formed by stars. Moreover, the material that formed into M31 satellites would necessarily have come from the outer parts of the M31 disk, which is generally less enriched (e.g. Gregersen et al. 2015, figure 9). For all these reasons, it is quite feasible that there would be no observable difference between the chemical abundances of M31 satellites even if they had very different formation scenarios (Kroupa 2015). It will be interesting to see if some difference is eventually found, although this might be much easier around our Galaxy than around M31. This could take advantage of Sextans not being part of the MW satellite plane (Casetti-Dinescu et al. 2018).

Similarly to the MW and M31, the satellite system of Centaurus A (Cen A) is also highly flattened and co-rotating, as evidenced by a radial velocity gradient across it (Müller et al. 2018). Such structures thus appear to be common, a claim also made by Ibata et al. (2014a) based on their finding that satellites on either side of a host galaxy have radial velocities of opposite signs, once the host systemic motion is accounted for. This is too recent for the debate to have settled (Cautun et al. 2015; Ibata et al. 2015). Nonetheless, it does seem like the Universe may well be full of TDGs if even just a few form in each galactic interaction (Okazaki & Taniguchi 2000). A high frequency of TDGs is also suggested by the correlation between cases where their existence is confirmed and the bulge mass fraction of the central galaxies (López-Corredoira & Kroupa 2016). If TDGs are more common, it would be easier to test whether the acceleration discrepancy persists in such systems, potentially resolving the question of how it arises in general.

2 OVERVIEW OF THE PORTFOLIO

To investigate the cause of the acceleration discrepancy, this portfolio considers several tests of the ΛCDM paradigm and some tests of MOND. The first of these (Banik & Zhao 2015a) is described in Section 3 and relates to the rather high collision velocity of the components of the Bullet Cluster, two interacting galaxy clusters (Tucker et al. 1995). Such a high velocity appears difficult to reconcile with ΛCDM (Thompson & Nagamine 2012; Kraljic & Sarkar 2015). However, the relative velocity between the clusters is mostly in the plane of the sky. Thus, it has not been directly measured but only estimated based on hydrodynamic simulations attempting to reproduce observed properties of the Bullet Cluster (Lage & Farrar 2014).

Fortunately, Molnar et al. (2013a) showed that it should soon become possible to measure the proper motion of the components of this cluster using the Moving Cluster Effect (MCE, Birkinshaw & Gull 1983). The MCE involves measuring redshifts of a background object multiply imaged by a foreground lens. Motion of the lens makes its potential time-dependent, thus giving the images different redshifts. However, the images could also have different redshifts because they have different magnification patterns across the source, provided this has a redshift gradient e.g. due to rotation. Although the issue could be resolved by taking integral field unit spectra at the appropriate velocity resolution, this is extremely challenging − only a spatially unresolved spectrum of each image is likely to be available for the foreseeable future.

Thus, I considered how these different effects could be disentangled using spectral line profiles of the individually unresolved images (Banik & Zhao 2015a). I also considered observational strategies to minimise the effects of such systematic errors, thus clarifying the kinematics of the Bullet Cluster. The same techniques could be applied to other interacting galaxy clusters like El Gordo (ACT-CL J0102-4915), which may be particularly problematic for ΛCDM due to its combination of high redshift ($z = 0.87$, Menanteau et al. 2012), high mass (Jee et al. 2014) and high inferred collision speed (Molnar & Broadhurst 2015).

Although relative proper motions may eventually be obtained in such systems, full 3D position and velocity information is only available within the LG out to about the distance of M31 (van der Marel et al. 2012b) and M33 (Brunthaler et al. 2005). Thus, the remainder of this portfolio focuses on the LG. The second work in the portfolio (Banik & Zhao 2016) − described in Section 4 − describes the construction of an axisymmetric dynamical model of the LG in ΛCDM, building on earlier spherically symmetric models (Kahn & Woltjer 1959; Sandage 1986; Peñarrubia et al. 2014). An axisymmetric model is expected to be rather accurate due to the very small MW-M31 tangential velocity (van der Marel et al. 2012b) and the close alignment of Cen A with the MW-M31 line (Ma et al. 1998). This model is used to perform a timing argument analysis i.e. see if some combination of model parameters can match the observed positions and radial velocities of M31 and LG dwarf galaxies using cosmological initial conditions ($\boldsymbol{v}_{pec} = \boldsymbol{0}$ at early times). Despite a reasonable allowance for observational uncertainties and inaccuracies in my model as a representation of ΛCDM, a full grid search through the model parameters did not yield a model matching the observed kinematics of the LG. This is because some galaxies have very high radial velocities.

To investigate this issue further, I used a 3D model of the LG in ΛCDM (Banik & Zhao 2017), the third work in this portfolio (Section 5). It is based on a FORTRAN algorithm borrowed from P. J. E. Peebles (Peebles et al. 2011). Despite using a different code written by a different author in a different programming language, my results still indicated that several LG galaxies have much higher radial velocities than expected in ΛCDM. The typical discrepancy between observations and the best-fitting 3D model was actually slightly higher than in the 2D case, even though the 3D model includes the major mass concentrations within ~ 10 Mpc.

The fourth work in this portfolio (Banik & Zhao 2018b) is described in Section 6 and looks at these high-velocity galaxies (HVGs) in more detail. As part of this, I visited Peebles at Princeton in order to revisit the work of Section 5 by performing a more thorough search for the best-fitting 3D model. This only slightly improved the fit to observations, still leaving several HVGs. A similar conclusion was also

reached by Peebles (2017). Confident that the HVGs are real, Banik & Zhao (2018b) shows that they preferentially lie very close to a well-defined plane which passes close to both the MW and M31. In this work, I use a restricted N-body model of the LG in MOND to argue that such a HVG plane is a natural consequence of a past MW-M31 flyby. Several ΛCDM-based explanations for the observations are also considered, but none of them seem plausible.

As a result, the fifth and final work in this portfolio (Section 7) considers MOND in more detail, in particular how it works in our own Galaxy (Banik & Zhao 2018a). Although MOND has often been tested using rotation curves of galaxies (e.g. Famaey & McGaugh 2012), I focus on comparing it to the recently measured Galactic escape velocity curve over distances of 8−50 kpc (Williams et al. 2017a). Both its amplitude and radial gradient are well matched in a MOND Galactic model that also accounts for its rotation curve. In future, the constraints should tighten considerably with GAIA data (Perryman et al. 2001), much of which is expected to be released in April 2018.

In Section 8, I suggest future avenues of investigation and give my conclusions in Section 9. Despite MOND not being a complete theory, it is well-defined and highly predictive in a wide range of circumstances. Therefore, it should be directly testable in the near future.

3 EFFECTS OF LENS MOTION & UNEVEN MAGNIFICATION ON IMAGE SPECTRA (Banik & Zhao 2015a)

On a large scale, the collision speed distribution of interacting galaxy clusters can be quite sensitive to the underlying law of gravitation (Cai et al. 2014). Indeed, the high collision speed of the components of the Bullet Cluster (1E0657-56, Tucker et al. 1995) has been argued in favour of modified gravity (Katz et al. 2013). However, this speed is not directly measured as the collision is mostly in the plane of the sky. Instead, the speed is estimated using simulations of the shock generated in the gas by the collision (Lage & Farrar 2014). The separation of the DM and gas (Clowe et al. 2006) also plays an important role − there is less gas drag at lower speeds, reducing the separation.

A collision speed close to 3000 km/s is considered necessary to explain the observed properties of the Bullet Cluster (Mastropietro & Burkert 2008). For the inferred masses of its components (Clowe et al. 2004), this appears difficult to reconcile with ΛCDM (Thompson & Nagamine 2012). This work suggested that a cosmological simulation requires a co-moving volume of $(4.48h^{-1}\mathrm{Gpc})^3$ to see an analogue to the Bullet Cluster. A subsequent analysis also found that systems analogous to the Bullet Cluster are expected to be rare in ΛCDM (Kraljic & Sarkar 2015).

Moreover, a few other massive colliding clusters with high infall velocities have been discovered in the last few years (Gómez et al. 2012; Menanteau et al. 2012; Molnar et al. 2013b). The El Gordo Cluster (ACT-CL J0102-4915) may be particularly problematic due to its combination of high redshift ($z = 0.87$, Menanteau et al. 2012), high mass (Jee et al. 2014) and high inferred collision speed (Molnar & Broadhurst 2015).

Molnar et al. (2013b) argue that inferring collision speeds from observations of the shock can be non-trivial just due to projection effects, let alone other complexities of baryonic physics. To see if there is any tension with ΛCDM, collision speeds should be determined in a more direct way. This is normally achieved using proper motions, but obtaining them is not feasible over cosmological distances.

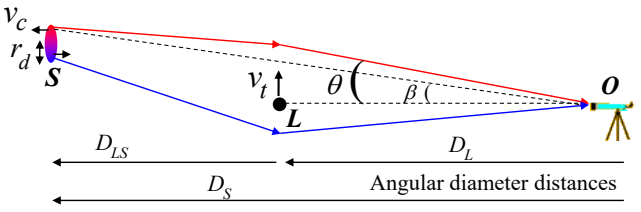


Figure 4. The lensing geometry is depicted here. Upper photon trajectory = primary image (same side as unlensed source), lower trajectory = secondary image. Relevant angular diameter distances are indicated at bottom. The lens L is treated as a point mass moving transversely to the viewing direction at speed v_t. The source S is an extended disk galaxy with scale length r_d. There is a redshift gradient across it due to rotation at speed $v_c\,(r)$, where r is distance from the source galaxy's centre.

Fortunately, the tangential motion of a massive object can be constrained using the Moving Cluster Effect (Birkinshaw & Gull 1983). The MCE relies on the gravitational potential of an object being time-dependent due to its motion. Consequently, if a source behind the object were multiply imaged, the images would have slightly different redshifts. Moreover, as DM generally outweighs gas on cluster scales (Blaksley & Bonamente 2010), the MCE is mostly sensitive to motion of the DM. This is simpler to model than gas, making the results easier to compare with cosmological simulations.

3.1 Method

The basic geometry is shown in Figure 4. To understand the MCE, consider a static universe in which the observer and source have negligible peculiar motions compared to the lens, which has a transverse velocity $\boldsymbol{v}_t$ in addition to any line of sight velocity that is not relevant for this analysis. It helps greatly to transform reference frame to the one in which the lens is static but the observer and source are moving at $-\boldsymbol{v}_t$. The lensing potential is now static. Emitted and received photons have different frequencies because the photon trajectories are not orthogonal to the velocities of the observer or the source. The difference in image redshifts can be expressed as a velocity shift

$$\Delta v_r|_{MCE} = -\boldsymbol{v}_t\cdot(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2) \tag{12}$$

Here, the light deflection angle for each image i is $\boldsymbol{\alpha}_i$. Using the thin-lens approximation[1], this can be expressed in terms of the observed image positions $\boldsymbol{\theta}_i$ as

$$\Delta v_r|_{MCE} = \frac{D_s}{D_{ls}}\boldsymbol{v}_t\cdot(\boldsymbol{\theta}_1-\boldsymbol{\theta}_2) \tag{13}$$

The angular diameter distances relevant to this problem

[1] i.e. the deflection occurs over a very small fraction of the entire photon path

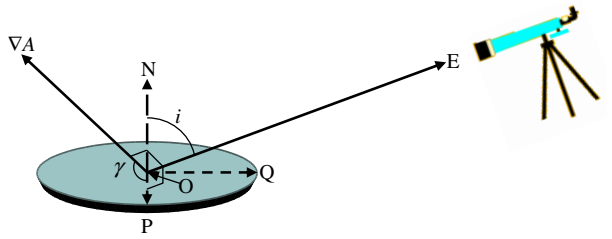


Figure 5. The observing geometry is shown here. The source galaxy has centre O and normal to its plane $\boldsymbol{ON}$. Earth is towards $\boldsymbol{OE}$, so the galaxy's inclination to the sky plane is i. $\boldsymbol{OQ}$ and $\boldsymbol{OP}$ are in the galaxy's plane and orthogonal to each other, with $\boldsymbol{OQ}$ as closely aligned with $\boldsymbol{OE}$ as possible. Thus, $\boldsymbol{OP}$ and $\boldsymbol{OE}$ are orthogonal. $\boldsymbol{\nabla} A$ is directed within the source plane, so must also be orthogonal to $\boldsymbol{OE}$. $\boldsymbol{\nabla} A$ is at an angle γ to $\boldsymbol{OP}$. The source is parametrised using cylindrical polar co-ordinates $(r,\ \phi)$, with centre O and initial direction $(\phi = 0)$ along $\boldsymbol{OQ}$.

are illustrated in Figure 4, with D_{ls} representing the angular diameter distance to the lens as perceived by an observer at the source when the photons we detect now most closely approached the lens. A source perfectly aligned with a point-like lens of mass M would appear as an Einstein ring of angular radius

$$\theta_E \ \equiv \ \sqrt{\frac{4GM}{c^2}\frac{D_{ls}}{D_l D_s}} \tag{14}$$

Combined with the source and lens positions, the lens mass M thus sets a typical angular scale for the problem. I use it to define

$$u \equiv \frac{\beta}{\theta_E} \qquad \text{and} \qquad y \equiv \frac{\theta}{\theta_E} \tag{15}$$

In terms of the lens and source physical properties and their true (unlensed) positions,

$$\Delta v_r|_{MCE} \ = \ \frac{2v_t\sqrt{GM\,(u^2+4)}}{c}\sqrt{\frac{D_s}{D_{ls}D_l}} \tag{16}$$

v_t is the component of the lens transverse velocity $\boldsymbol{v}_t$ directed along the separation between the observed images.

I estimated the effects of source & observer peculiar velocities – they should not affect Δv_r much. There is also a time delay between the images, causing us to observe the source at an earlier epoch in one image than in the other. Due to cosmic expansion, this creates a redshift difference, but only a very small one (< 1 m/s).

The main systematic uncertainty in MCE measurements is likely to be the Differential Magnification Effect (DME), an observational artefact due to our inability to take highly accurate spectra of the images while also spatially resolving them. This causes parts of the source with different redshifts to get blended together in spectra. The precise way in which this blending occurs differs between the images.

To model how such single-pixel spectra might look, I modelled the source as a typical spiral galaxy with exponential surface density profile (Freeman 1970) and a realistic rotation curve based on Equation 8. The lens is treated as a point mass. The parameters considered (Table 1) are designed for the Bullet Cluster (Tucker et al. 1995).

The basic idea is that spatially unresolved spectra can determine the intensity-weighted mean redshift $\overline{v_r}$ of each image. This may be affected by rotation of the source galaxy. The effect isn't reliant on an expanding Universe. Neglecting cosmic expansion for the moment, the mean redshift velocity of each image is

$$\overline{v_r} \ \equiv \ \frac{\int_{\text{Image}} A\Sigma v_r \ dS}{\int_{\text{Image}} A\Sigma \ dS} \tag{17}$$

The integrals are over area elements of the source S, which I treat as an exponential disk with surface density profile

$$\Sigma \ = \ \Sigma_0 \ e^{-\tilde{r}} \quad \text{where } \tilde{r} \equiv \frac{r}{r_d} \tag{18}$$

The magnification A varies little over the source galaxy as $\frac{r_d}{D_s} \ll \theta_E$ (Table 1). Thus, a linear approximation to A is sufficient.

$$A \ \approx \ A_0 + \frac{\partial A}{\partial u}du \quad (A_0 \equiv A \text{ at centre of source}) \tag{19}$$

The mean redshift of each image due to the DME is

$$|\overline{v_r}| = \tag{20}$$

$$\frac{v_f r_d \sin i \cos\gamma}{D_s\theta_E}\frac{\int_0^{\infty}\int_0^{2\pi}\overbrace{e^{-\tilde{r}}}^{\propto\Sigma}\widetilde{v_c}(\tilde{r})\tilde{r}^2\overbrace{\frac{4}{u^2(u^2+4)^{\frac{3}{2}}}}^{-\frac{\partial A}{\partial u}}\sin^2\phi\ d\phi\ d\tilde{r}}{\underbrace{\pi\left(\frac{u^2+2}{u\sqrt{u^2+4}}\pm 1\right)}_{\propto A}}$$

I obtained a family of rotation curves using the 'simple μ-function' in MOND (Famaey & Binney 2005), as discussed just before Equation 8. The rotation curve shape is determined by the central surface density, which I parametrise using

$$k \ \equiv \ \frac{\Sigma_0 G}{a_0} \tag{21}$$

Thus, my rotation curves flatline at

$$v_f \ = \ \sqrt[4]{GMa_0} \tag{22}$$

$$= \ \sqrt[4]{2\pi k}\sqrt{r_d a_0} \tag{23}$$

After making a few approximations to estimate $\boldsymbol{g}_N$, I obtained the overall rotation curve shape

$$\widetilde{v_c}(\tilde{r}) \equiv \frac{v_c(\tilde{r})}{v_f} = \frac{\sqrt{\pi k\tilde{r}f(\tilde{r}) + \tilde{r}\sqrt{(\pi k f(\tilde{r})+1)^2 - 1}}}{\sqrt[4]{2\pi k}} \tag{24}$$

$$f(\tilde{r}) = \frac{1-\frac{13}{4}e^{-\tilde{r}}}{\tilde{r}^2} - \frac{7e^{-\tilde{r}}}{4\tilde{r}} + \frac{9\left(1-e^{-\tilde{r}}-\tilde{r}e^{-\tilde{r}}\right)}{2\tilde{r}^4} \tag{25}$$

The normalised rotation speed at radius $\tilde{r} \equiv \frac{r}{r_d}$ is $\widetilde{v_c} \equiv \frac{v_c}{v_f}$, with v_f the flatline level of the source galaxy rotation curve. Two example rotation curves are shown in the top panel of Figure 6. Its bottom panel shows the ratio between v_{max} and v_f for disks with different central surface densities.

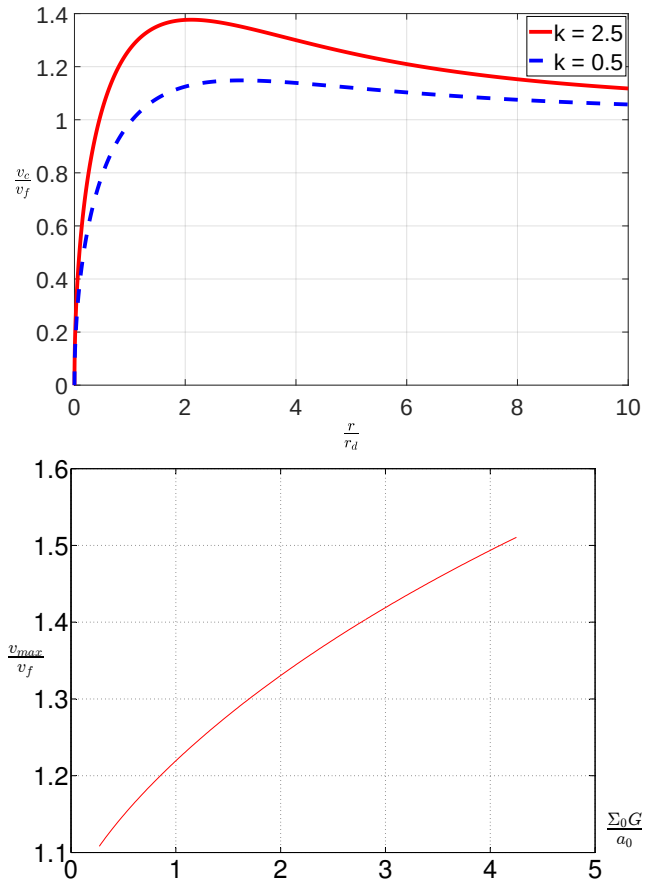


Figure 6. *Top*: Rotation curves resulting from Equation 24. $v_c\,(r)$ flatlines at v_f. The surface density $\Sigma = \Sigma_0 e^{-\frac{r}{r_d}}$. The parameter $k \equiv \frac{G\Sigma_0}{a_0}$ controls the shape of the rotation curve. *Bottom*: The ratio of maximum to flatline rotation speed as a function of k.

3.2 Results

Combining Equation 20 with the rotation curve shape from Equation 24, I get that

$$\Delta\overline{v_r}|_{DME} = \frac{v_f\ r_d\ \sin i\ \cos\gamma\ I\ c\ \sqrt{D_l}}{\sqrt{u^2+4}\ \sqrt{GMD_{ls}D_s}} \quad (26)$$

$$I \equiv \int_0^\infty \mathrm{e}^{-\tilde{r}}\tilde{v_c}(\tilde{r})\ \tilde{r}^2 d\tilde{r} \quad (27)$$

The integral I depends on the central surface density k. However, the ratio $\frac{I}{\tilde{v}_{max}} = 1.89 \pm 0.02$ over the range $k = 0.1 - 5$. Thus, k is not needed very precisely if v_{max} were available rather than v_f. In practice, it is much easier to obtain v_{max}.

Using the parameters in Table 1, I obtained the results shown in Figure 7. The MCE and DME are comparable if the source is a typical spiral galaxy. This suggests that the DME might well confuse measurements of the MCE without certain precautions. Unfortunately, it can be difficult to calculate the DME and adjust for it because it relies on quantities that may be difficult to determine e.g. the variation of A with position and which side of the galaxy is the approaching side. Thus, I considered whether the detailed profiles of individual spectral lines could be used to distinguish the DME from the MCE. A detailed line profile would contain much more information than just the centroid location $\overline{v_r}$.

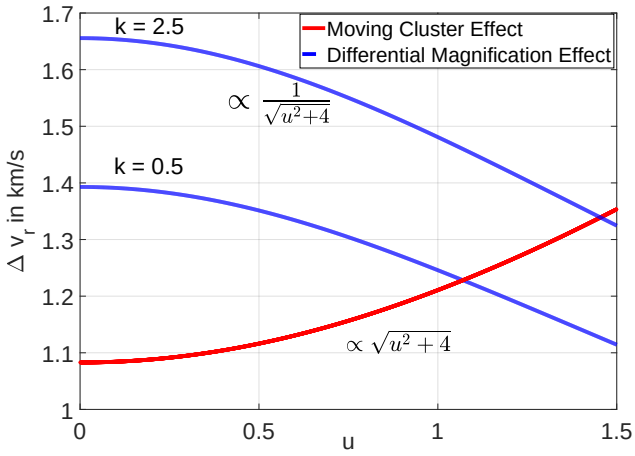


Figure 7. The difference in redshift between double images of a typical background galaxy as a function of its position, due to the effects described in the text (Equations 16 and 26). Parameter values used here are listed in Table 1. The DME is affected by the shape of the rotation curve (governed by the central surface density k) once its flatline level v_f is fixed. If instead v_{max} is known, then k has only a very small ($\sim 1\%$) impact on the DME.

Table 1. Parameters used for Figure 7. The source galaxy is assumed positioned so as to maximise the MCE (i.e. it is separated from the lens on the sky along the direction of motion of the lens, which is clear from images). The lens mass should roughly correspond to the sub-cluster in the Bullet. A flat ΛCDM cosmology is adopted (Planck Collaboration XVI 2014).

Parameter	Meaning	Value
H_0	Present Hubble constant	67.3 km/s/Mpc
Ω_m	Present matter density	0.315
D_l	(Angular diameter) distance to lens at $z_l = 0.296$	0.945 Gpc
D_s	Distance to source at $z_s = 1.7$	1.795 Gpc
D_{ls}	Distance to source from lens position in spacetime	1.341 Gpc
M	Mass of lens	$1.2 \times 10^{14} M_\odot$
r_d	Scale length of source galaxy	3.068 kpc
v_t	Tangential velocity of lens	3000 km/s
v_f	Flatline level of source galaxy rotation curve	100 km/s
$\sin i \cos\gamma$	See Figure 5. Isotropic average.	$\frac{1}{2}$

I began by determining if existing observatories could attain the required spectral resolution within a reasonable timeframe. To this end, I considered the Atacama Large Millimetre Array (ALMA). Using the online calculator, I found that ALMA probably can resolve individual spectral lines well enough to distinguish the MCE from the DME (Table 2).

To take advantage of this, I performed calculations to see how the DME and MCE affect individual line profiles. I began by mapping the radial velocity of an edge-on disk galaxy.

$$v_r\,(r,\phi) = v_c\,(r)\sin\phi \quad (28)$$

Table 2. Input parameters used for the ALMA exposure time calculator, available at:

https://almascience.eso.org/proposing/sensitivity-calculator

The dual polarisation mode should be used as polarisation is unimportant here. The angular resolution does not affect the result, which was 6.17 hours.

Parameter	Value
Declination	$-56°$
Frequency	150 GHz
Bandwidth per polarisation	100 m/s
Water vapour column density	5^{th} octile (1.796 mm)
Number of antennas	50 × 12 metre
Root mean square (rms) sensitivity	1.5 mJy

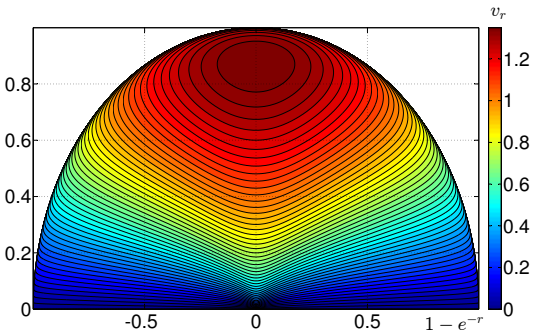


Figure 8. Radial velocity map of a disk galaxy viewed by an observer within its plane at large x (far off to the right), for the case $k = 2.5$ (similar to the MW). Radial velocities are antisymmetric about the x-axis. The radial co-ordinate has been rescaled so the displayed size of each region is proportional to its brightness. The units are such that $r_d = 1$ and $v_f = 1$. Note the large region with v_r close to its maximum value. The result for $k = 0.5$ is very similar, although v_{max} is much closer to v_f (Figure 6).

The results are shown in Figure 8 for the case $k = 2.5$, similar to the MW value. I then binned the galaxy in r and ϕ in order to take advantage of Equation 28 being separable in (r, ϕ). The key trick is to calculate v_c only once at each r, for all ϕ. In this way, I determined v_r and thereby classified the luminosity of the galaxy according to the radial velocity of the region emitting the light. This allows a synthetic line profile to be constructed. Four examples are shown in Figure 9, three of which allow for some random motion as well as ordered circular motion. The line profiles show a distinct horn corresponding to the 'bull's-eye' towards the top of Figure 8. This feature is due to a turning point in the rotation curve, causing a large part of the galaxy to have a similar v_r.

To allow for the MCE, I simply translated the line profile. For the DME, I let

$$A = 1 + n\tilde{r}\sin\phi \tag{29}$$

For reference, I constructed a control line profile with

$$A = 1 \; \forall_{r,\phi} \tag{30}$$

This control line profile was subtracted from the profiles modified by the DME and the MCE. The residuals are shown in Figure 10 for $n \ll 1$ but large enough to avoid numerical issues. The pattern of residuals is quite different in the two

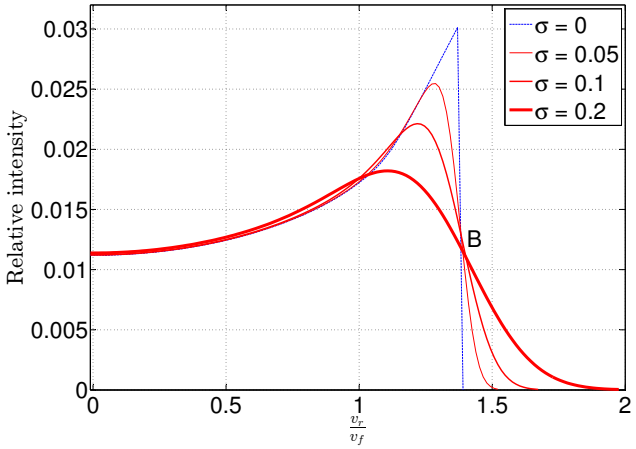


Figure 9. The synthetic line profile of an intrinsically narrow spectral line in an unlensed galaxy with $k = 2.5$, viewed edge-on. The profile is symmetric about $v_r = 0$. The sharp drop in the line profile (blue) would probably get blurred (e.g. by random motions), so I convolved the profile with Gaussians of width σ (given in units of the flatline rotation curve level v_f). The results are shown as red lines with thickness $\propto \sigma$. Notice how all 4 profiles pass close to the point marked B. The result for $k = 0.5$ is similar, if velocities are scaled to v_{max} rather than v_f.

cases, even though both cause the same $\Delta\overline{v_r}$. This might well allow the MCE and DME to be distinguished. In particular, the MCE can lead to very large residuals close to the peaks in the line profile, depending on how sharp they are. Even the less dramatic features near $v_r = 0$ might be useful – the MCE and DME give opposite signs for the residuals in this spectral region despite causing the same overall redshift difference between the images.

3.3 Observational strategies

The MCE can be enhanced relative to the DME by a number of strategies, especially if there is a choice of which multiple images to target for detailed spectroscopic follow-up. Avoiding a spiral galaxy as a target reduces the DME, although an elliptical can still rotate. Even a low-resolution spectrum should be able to distinguish a fast-rotating spiral from an elliptical with mild rotation in the region emitting most of the light. This is because spirals ought to have a characteristic double-horned spectral profile whereas ellipticals would have a roughly Gaussian profile. However, the latter lacks sharp features, making the MCE itself harder to detect by raising random errors.

Spiral galaxies are acceptable targets if they are viewed face-on as such objects have little gradient in v_r across them. However, even an edge-on spiral can make a good target if it is oriented so its major axis is a direction along which A hardly changes. For a point mass lens, this would mean the source galaxy's major axis was orthogonal to the apparent lens-source line. In a more complicated lens, it might be possible to estimate how A varies with sky position and use this magnification map to guide the selection of targets.

Some interesting possibilities arise if the source is very inhomogeneous. A small region might be forming stars

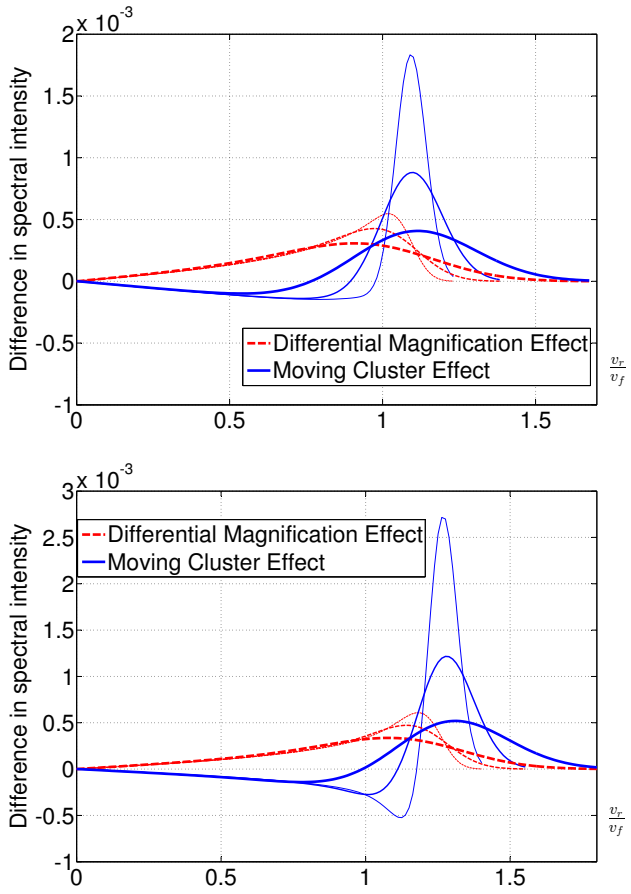


Figure 10. The residuals in the spectral profile due to the DME (Equation 29) and the MCE (horizontal shift of profile), obtained by subtracting a control line profile (Equation 30). The patterns are antisymmetric about $v_r = 0$. Results are shown for an edge-on galaxy with $k = 0.5$ (*top*) and $k = 2.5$ (*bottom*). Both effects change the mean redshift by 1% of the maximum rotation speed, representing 1.08% of v_f for $k = 0.5$ and 1.26% for $k = 2.5$. The spectra were convolved with Gaussians of widths 0.05, 0.1 and 0.2 v_f (higher σ indicated by thicker line). The MCE can't change the amplitudes of the horns while the DME can − it makes one more pronounced and the other less.

rapidly and emit strongly in the far-infrared due to dust. Targeting only spectral lines at these wavelengths then reduces the DME because the emitting region is small and A varies only a little over it. However, the MCE is unaffected by the size of the emitting region (as long as it is much smaller than the Einstein radius).

One can target fainter spiral galaxies so the source is likely to be smaller and slower-rotating. This strategy may be difficult to implement with current technology. It is promising in the long run because there are many more fainter galaxies than brighter ones (Schechter 1976).

I expect that the MCE can be measured in the near future if careful consideration is put into reducing the impact of the DME. Important insights may be gained by comparing the detailed line profiles between different images of the same object. If a target was used for which the DME should be negligible, then the validity of any claimed MCE detection could be checked by comparison of the observed pattern of residuals between appropriately scaled spectra of the multiple images. If instead the DME is not negligible, its magnitude could be estimated from this pattern of residuals because of the very different ways in which differential magnification and lens motion affect spectral line profiles, even if they cause an equal difference between the mean image redshifts (Figure 10). A promising target for detecting the MCE might be the triply imaged galaxy discovered by Gonzalez et al. (2009).

4 DYNAMICS OF THE LOCAL GROUP IN ΛCDM (Banik & Zhao 2016)

Section 3 discussed possible tests of gravity based on galaxy clusters, where the collision velocity may be several hundred km/s faster in MOND than would be feasible in ΛCDM (Katz et al. 2013, figure 8). Even so, it is difficult to actually perform this test because of the cosmological distances to these interesting systems (e.g. the Bullet Cluster has a redshift of $z = 0.296$, Tucker et al. 1998). This makes it difficult to know the 3D position and velocity structure of the system.

Although gravity could be tested without all 6 phase space co-ordinates, it is at least necessary to know the peculiar velocities. As velocity errors are relatively small nowadays, the dominant uncertainty arises from redshift-independent distances. Assuming a 10% distance uncertainty and that peculiar velocities might differ by $\Delta v = 100$ km/s between the different gravity theories, this suggests that it is difficult to properly test MOND beyond a distance

$$d_{max} = \frac{\Delta v}{0.1 H_0} \quad (31)$$

$$= 15 \text{ Mpc} \quad (32)$$

A good test of MOND should involve a system where its prediction differs by $\sim 5\sigma$ from that of ΛCDM. This would reduce d_{max} to only 3 Mpc. Therefore, the remainder of this portfolio will focus on the LG.

A major alteration to the gravitational field would undoubtedly have far-reaching implications for the motions of LG galaxies. In particular, MOND implies there was a past close MW-M31 flyby (Zhao et al. 2013) at a much faster relative velocity than expected in ΛCDM, a model in which such a flyby is precluded due to dynamical friction between DM halos (Privon et al. 2013). These high velocities would allow the MW and M31 to gravitationally slingshot any passing LG dwarf galaxies out at high speeds in 3-body interactions (MW, M31 and dwarf). In this section, I summarise the work of Banik & Zhao (2016) where I investigated whether there is any evidence for such a scenario using an axisymmetric dynamical model of the LG in ΛCDM.

The basic idea behind this is called the timing argument (Kahn & Woltjer 1959). It involves using observed peculiar velocities $\boldsymbol{v}_{pec}$ to estimate the gravitational field $\boldsymbol{g}\,(t)$ within the LG throughout cosmic history, exploiting the fact that $\boldsymbol{v}_{pec}$ was rather small early in the history of the Universe (Planck Collaboration XXVII 2014). Similar analyses were done previously by Sandage (1986) and by Peñarrubia et al. (2014), the latter using a spherically symmetric model of the LG to estimate how the MW and M31 should slow the out-

ward recession of LG dwarfs in ΛCDM. Consideration of 12 LG analogues in cosmological ΛCDM simulations (Fattahi et al. 2016) shows that such timing argument calculations are fairly accurate (Peñarrubia & Fattahi 2017).

To improve the accuracy further, I developed a more advanced axisymmetric model of the LG. Like Peñarrubia et al. (2014), I also assumed that the mass in the LG is entirely contained within the MW and M31, which I took to be on a radial orbit. Recent proper motion measurements of M31 indicate only a small tangential velocity relative to the MW, making the true orbit almost radial (van der Marel et al. 2012a). I also included Cen A in my models because it lies rather close to the MW-M31 line (Ma et al. 1998).

4.1 Method

I adopt a standard flat ΛCDM cosmology. Ignoring components other than matter and dark energy, the evolution of the cosmic scale factor $a\,(t)$ can be determined analytically and is fully specified by the present Hubble constant H_0 and matter density parameter $\Omega_{m,0}$.

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho_m - 2\rho_\Lambda\right) \quad (33)$$
$$= {H_0}^2\left(-\frac{1}{2}\Omega_{m,0}\ a^{-3} + \Omega_{\Lambda,0}\right). \text{ Therefore,}$$

$$a(t) = \left(\frac{\Omega_{m,0}}{\Omega_{\Lambda,0}}\right)^{\frac{1}{3}} \sinh^{\frac{2}{3}}\left(\frac{3}{2}\sqrt{\Omega_{\Lambda,0}}H_0 t\right) \quad (34)$$

To integrate test particle trajectories, it is necessary to first have trajectories for the MW, M31 and Cen A that match the presently observed distances of M31 and Cen A to within 1 kpc. I do this using a 2D Newton-Raphson algorithm on the initial relative positions of all three galaxies along a line.[1] Initial velocities were found using

$$\boldsymbol{v}_i = H_i \boldsymbol{r}_i \quad \text{where } H_i \equiv \frac{\dot{a}}{a} \quad \text{when } t = t_i \quad (35)$$

To minimise convergence issues, I used each solution to the problem for the next set of model parameters, keeping the parameter changes small. Figure 11 shows the MW-M31 separation in the massive object solution for one of my models.

With the trajectories of the massive galaxies in hand, I solved test particle trajectories starting at some early initial time t_i. I took the barycentre of the LG then as the centre of expansion. The initial velocities followed a pure Hubble flow (Equation 35) because the Universe was nearly homogeneous at early times − peculiar velocities on the last scattering surface are only ∼3 m/s (Planck Collaboration XIII 2016) whereas the present value is over 100 km/s for M31 (McConnachie 2012; van der Marel et al. 2012b).

Test particle trajectories were advanced using a fourth-order Runge-Kutta scheme with an adaptive but quantised timestep that was varied in powers of 2. Any close approaches to the MW or M31 were treated as an accretion event, increasing the mass of the accreting galaxy by an amount proportional to the volume represented by each test

[1] For stability, I under-relaxed the algorithm i.e. in each iteration, I altered the parameters by 80% of what the algorithm would normally have altered them by

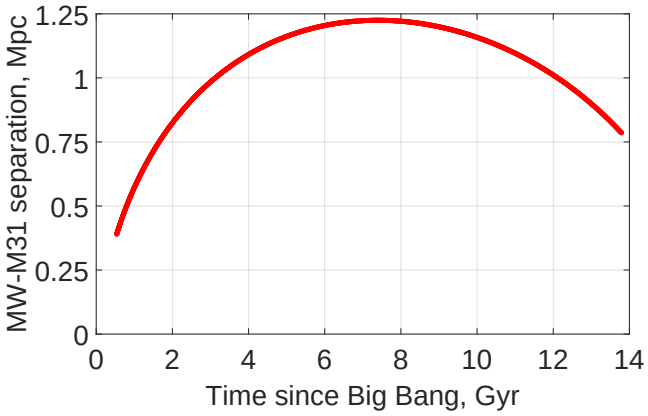


Figure 11. MW−M31 separation $d(t)$ for a typical model where $q_1 = 0.2$ and $M_i = 3.4 \times 10^{12} M_\odot$ (parameters defined in Table 3). $d(t)$ always looks broadly similar − in ΛCDM, the MW and M31 have never approached each other closely for any plausible model parameters.

particle. This required the test particle trajectories to be recalculated. I found that the MW and M31 masses converged very well after just a single iteration. The resulting velocity field for the LG is shown in the top panel of Figure 12. Its bottom panel shows the resulting LG Hubble diagram, comparing velocity directly away from the LG barycentre with distance from there.

To compare with observations, I obtained a test particle trajectory ending at the same position as a 'target' LG galaxy. The Galactocentric radial velocity (GRV) of the observed target and the simulated test particle were then compared to judge how well the model fits observations. My list of targets was almost the same as that used by Peñarrubia et al. (2014), with a few minor alterations. A wide range of plausible model parameters (Table 3) were investigated within the context of ΛCDM.

In a homogeneous Universe, $\boldsymbol{r} \propto a\,(t)$ so that $\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r}$. Including additional forces arising from the gravity of the MW, M31 and Cen A, this makes the equation of motion for each test particle

$$\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r} - \sum_{\substack{\text{j = MW,}\\ \text{M31, Cen A}}} \frac{GM_j\left(\boldsymbol{r} - \boldsymbol{r}_j\right)}{\left(|\boldsymbol{r} - \boldsymbol{r}_j|^2 + {r_{S,j}}^2\right)^{\frac{1}{2}}|\boldsymbol{r} - \boldsymbol{r}_j|^2} \quad (36)$$

The force towards each massive galaxy is not $\propto \frac{1}{r^2}$ at low distance r from it because this would contradict observed flat rotation curves of major LG galaxies (e.g. Carignan et al. 2006). To be consistent with the observed values of v_f for the MW and M31 (180 and 225 km/s, respectively); I had to soften the force below a radius $r_S = \frac{GM}{{v_f}^2}$. This radius is different for the MW and M31. For Cen A, I used $r_S = 100$ kpc because this analysis does not consider LG particles ending up near Cen A, making the precise force law used there irrelevant.

Test particles were started on a grid of plane polar coordinates. I assumed that the initial masses of the MW and M31 arose by completely depleting all the mass in some region, so I did not start any test particles within it. Assuming

the 'feeding zone' to be delimited by an equipotential, I first determined the potential resulting from the MW and M31.

$$U = \sum_{j=MW,M31} -\frac{GM}{r_{S,j}} Ln\left(\frac{\sqrt{1+{b_j}^2}-1}{b_j}\right) \quad (37)$$

$$b_j \equiv \frac{|\boldsymbol{r}-\boldsymbol{r_j}|}{r_{S,j}} \quad (38)$$

Next, I determined the volume V of the accretion region such that $\rho_M V = M_i$, where M_i is the initial combined mass of the MW & M31 in my model while ρ_M is the cosmic mean matter density at t_i.[1] Finally, I determined the equipotential U_{exc} such that the region with $U < U_{exc}$ has volume V. Test particle trajectories were not started in the region where $U < U_{exc}$ at $t = t_i$.

To get target galaxies in the same co-ordinate system as used in my simulation (which has its $y-$axis aligned with the MW-M31 separation), I used

$$x = d_{MW}\,|\hat{\boldsymbol{d}}_{MW} \times \hat{\boldsymbol{r}}_{MW}| \quad (39)$$

$$y_{rel} \equiv y - y_{MW} = d_{MW}\left(\hat{\boldsymbol{d}}_{MW} \cdot \hat{\boldsymbol{r}}_{MW}\right) \quad (40)$$

The vector from the MW to a target galaxy is denoted $\boldsymbol{d}_{MW}$ while the direction from M31 towards the MW is denoted $\hat{\boldsymbol{r}}_{MW}$. I use the convention that $\hat{\boldsymbol{v}} \equiv \frac{\boldsymbol{v}}{v}$ for any vector $\boldsymbol{v}$ with length $v \equiv |\boldsymbol{v}|$.

4.2 Comparison with observations

Close to the MW and M31, the velocity field is complicated because there are intersecting trajectories (Figure 12). This makes it impossible to uniquely predict the velocity based on position. As a result, I had to exclude any target galaxies which fell in such regions. Fortunately, this was extremely rare, mainly because of pre-selection of targets by Peñarrubia et al. (2014).

The key aspect of my algorithm was carefully determining the model-predicted GRV of each target galaxy. To do this, I obtained a test particle trajectory landing at the same position as each target. I started with the test particle that landed nearest to its observed position. I then applied the 2D Newton-Raphson algorithm to the initial position of this test particle, updating its initial velocity according to Equation 35. I considered this process to have converged once the final position error fell below 0.001% of the target's distance from the LG barycentre. This trajectory was then used to determine the model-predicted GRV of the target galaxy.

$$GRV_{model} = \frac{v_x x + (v_y - \dot{y}_{MW})(y_{rel})}{\sqrt{x^2 + {y_{rel}}^2}} \quad (41)$$

To account for distance uncertainties, this procedure was repeated for each target with its heliocentric distance raised to the 1σ upper limit of the observed value. The difference between these GRV estimates is σ_{pos}, the uncertainty in the model-predicted GRV of a target due to uncertainty in its position along the line of sight.

To compare with observations, the observed GRV of each target is obtained by adjusting its observed heliocentric radial velocity (HRV) for the motion of the Sun within the MW. The solar motion $\boldsymbol{v}_\odot$ is mostly circular motion within the MW disk at speed $v_{c,\odot}$, the speed of a test particle on a circular Galactic orbit at the position of the Sun. It is useful to define a Local Standard of Rest (LSR), a reference frame rotating at this speed. Naturally, the Sun has a small amount of motion with respect to the LSR (magnitude and direction given in Table 3).

$$GRV_{obs} = HRV + \mathbf{v}_\odot \cdot \hat{\mathbf{d}}_{MW} \quad (42)$$

As well as the uncertainty on each GRV due to position σ_{pos} and error on the observed HRV σ_{v_h}, I added another term σ_{extra} to account for large scale structure, interactions between LG dwarf galaxies and other effects not included in my model.

$$\sigma = \sqrt{{\sigma_{pos}}^2 + {\sigma_{v_h}}^2 + {\sigma_{extra}}^2} \quad (43)$$

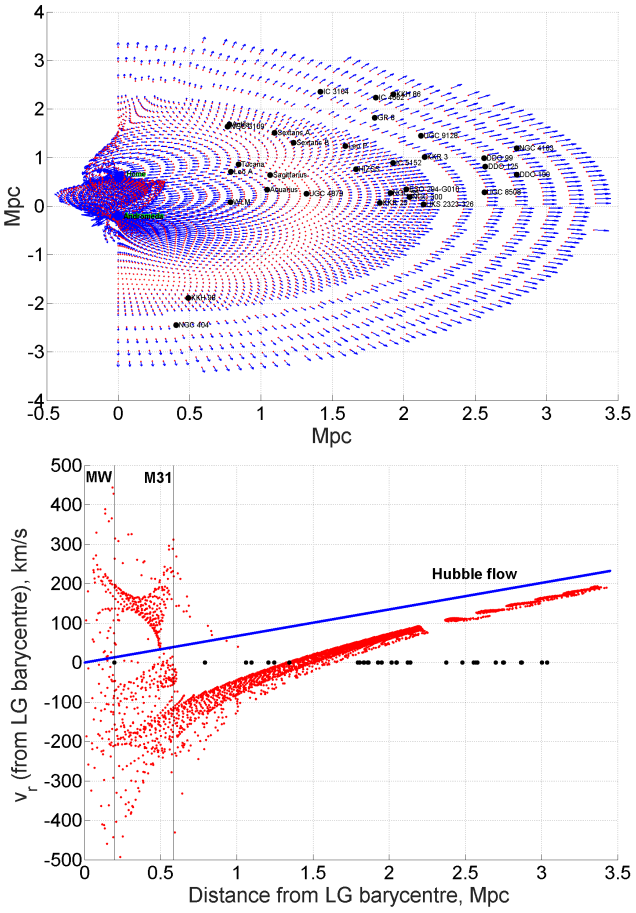


Figure 12. *Top:* Local Group velocity field for the case $q_1 = 0.3$, $M_i = 4 \times 10^{12} M_\odot$. Locations of target galaxies are overplotted as black dots with galaxy names given beside them. The MW is just above the centre. Only particles starting at $x > 0$ (and thus $v_x > 0$) were considered. Thus, the presence of particles at $x < 0$ indicates intersecting trajectories and a disturbed velocity field. *Bottom:* Radial velocities of test particles with respect to the LG barycentre. Vertical lines represent distances of the MW and M31 from there. Black dots indicate positions of target galaxies. Without proper motions, they can't be put on such a Hubble diagram at the correct velocity, so I show this as 0.

[1] ρ_M includes both baryons and DM.

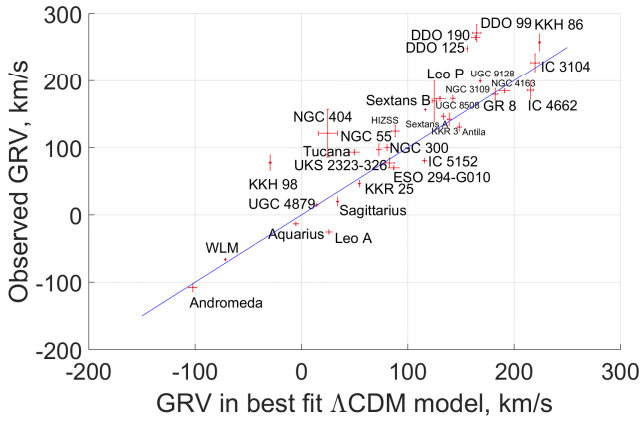


Figure 13. Comparison between model-predicted and observed Galactocentric radial velocities based on the most likely model parameters ($q_1 = 0.2$, $M_i = 3.4 \times 10^{12} M_\odot$). The line of equality is also shown in blue.

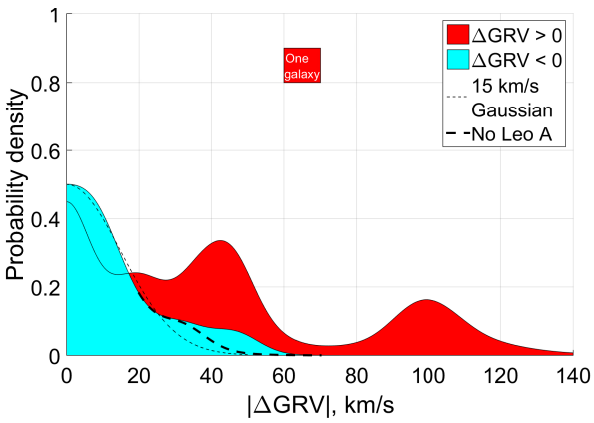


Figure 14. Histogram showing observed − predicted GRVs (i.e. ΔGRVs) of target galaxies using the most plausible model ($q_1 = 0.2$ instead of 0.14, other parameters as in Figure 13). Each data point was convolved with a Gaussian of width $\sigma = \sqrt{{\sigma_{pos}}^2 + {\sigma_{v_h}}^2 + {\sigma_{v_{c,\odot}}}^2}$. I divided the sample into targets with $\Delta GRV < 0$ (blue) and those with $\Delta GRV > 0$ (red). The area corresponding to one galaxy is shown as a red square. A Gaussian of width 15 km/s is overplotted as a short-dashed line. This matches the $\Delta GRV < 0$ subsample quite well, especially when Leo A is excluded (long-dashed line).

σ_{extra} should not exceed typical velocity dispersions/rotation speeds of LG dwarfs, which I estimate as ~ 15 km/s (e.g. Kirby et al. 2014). As the motion of M31 would be much harder to alter than the motion of a less massive LG dwarf, I kept σ_{extra} for M31 at $\frac{1}{10}$ of the value for other LG galaxies. However, the results are not much altered if M31 is treated in the same way as other LG dwarfs.

4.3 Results and discussion

The resulting probability distributions of the model parameters are shown in Figure 15 and summarised in Table 3. The posterior on σ_{extra} only allows values $\gtrsim 40$ km/s, which is rather high. I considered explanations such as a higher Hubble constant, altered start time, interactions with massive satellites like M33 and tides from objects outside the LG. None of these seem to work, though the last two possibilities are considered more thoroughly in Sections 5 and 6.

Additional inaccuracies in the model may arise from the effects of large scale structure and distant encounters between LG dwarfs. The likely magnitude of such effects can be estimated based on more detailed ΛCDM cosmological simulations. Considering LG analogues in such simulations, it has been found that the dispersion in radial velocity with respect to the LG barycentre at fixed distance from there should be $\sigma_H \sim 30$ km/s (Aragon-Calvo et al. 2011), insufficient to explain my high inferred σ_{extra}.

To better understand the discrepancy, I compare model-predicted and observed GRVs for the best-fitting model (Figures 13 and 14). These show that most velocities are more outwards than in the model, making it difficult to argue that the discrepancy arises from unmodelled interactions between LG dwarfs. Instead, I suggested that gravitational slingshot encounters with the MW or M31 flung out LG dwarfs at high speed. As this process is already accounted for, these massive galaxies must have been moving much faster than in my model. This would occur in MOND − their relative speed at closest approach would have been ~ 600 km/s (Bílek et al. 2017), very different to ΛCDM where it would rarely have been faster than the 110 km/s it is today (Figure 11).

Name	Meaning and units	Prior	Result
σ_{extra}	Extra velocity dispersion along line of sight, km/s	0 − 100	$45.1^{+7.0}_{-5.7}$
M_i	Initial MW + M31 mass, trillions of solar masses	2 − 6.6	4.1±0.3
q_1	Fraction of MW + M31 mass initially in the MW	0.04−0.96	0.14±0.07
$v_{c,\odot}$	Circular speed of MW at position of Sun, km/s	239 ± 5	239.5±4.8
	Fixed parameters		
d_0	Distance to M31, kpc	783 ± 25	
H_0	Hubble constant at the present time, km/s/Mpc	67.3	
$\Omega_{m,0}$	Present matter density in the Universe $\div \frac{3{H_0}^2}{8\pi G}$	0.315	
a_i	Scale factor of Universe at start of simulation	0.1	
$r_{acc,MW}$	Accretion radius of MW	15,337 parsecs	
$r_{acc,M31}$	Accretion radius of M31	21,472 parsecs	
$U_\odot$	Components of	14.1 km/s	
$V_\odot$	non-circular motion of	14.6 km/s	
$W_\odot$	Sun within Milky Way	6.9 km/s	

Table 3. Priors and 1σ confidence levels on model parameters. The latter are far from the boundaries imposed by the former, showing that the results are not strongly affected by prior assumptions. I used the measurement of d_0 by McConnachie (2012), cosmological parameters from Planck Collaboration XIII (2016), $v_{c,\odot}$ from McMillan (2011) and the Sun's non-circular velocity from Francis & Anderson (2014), given here in standard notation. Uncertainty in the latter is much less than in $v_{c,\odot}$, which was assumed to be within 3σ of its observed value.

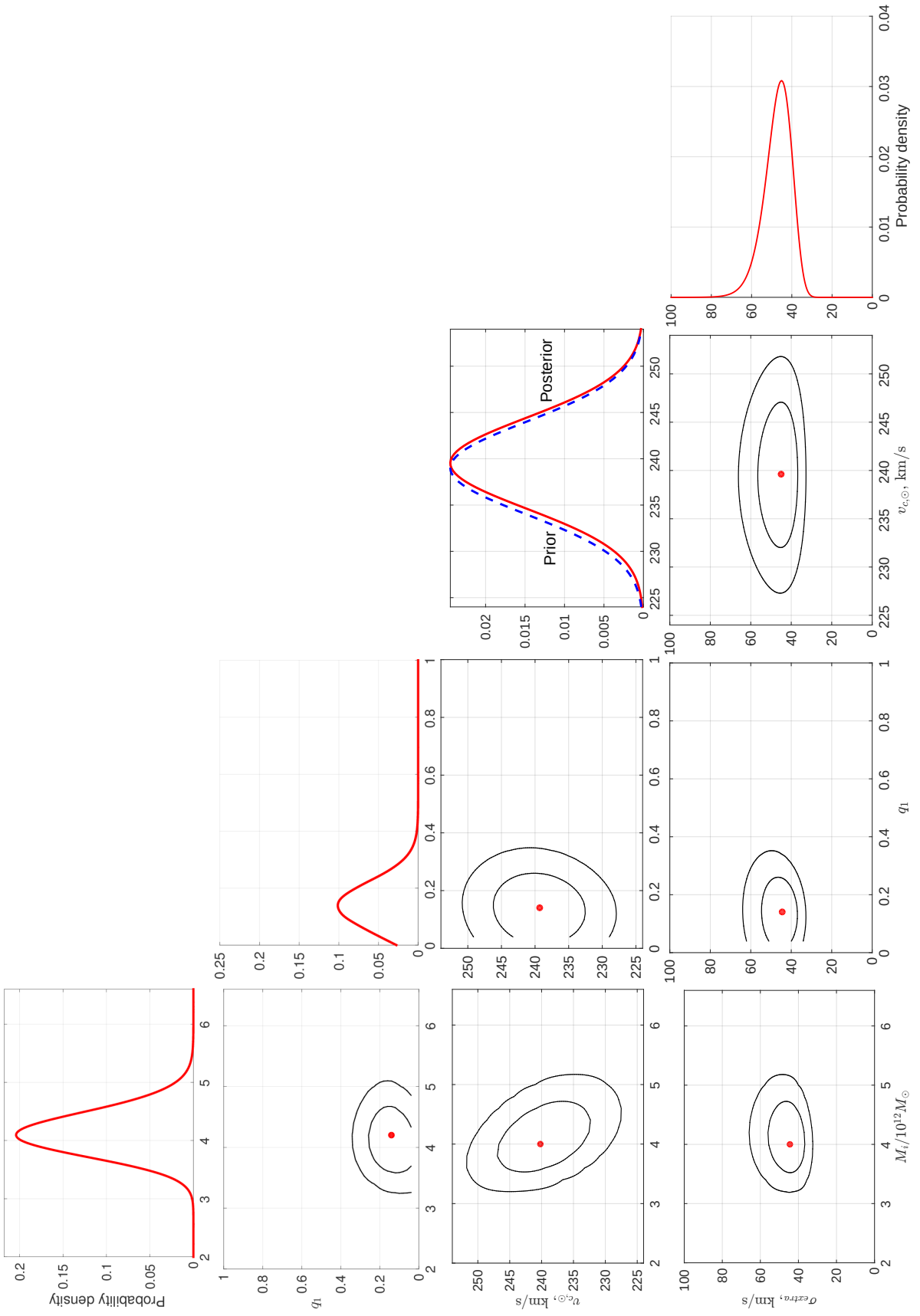


Figure 15. Marginalised posterior probability distributions of the free parameters defined in Table 3, with other parameters marginalised over. For variables plotted against other variables, I show the contours of the probability density which correspond to the usual 1σ and 2σ confidence levels, as well as the most likely pair of values. Rotate figure 90° clockwise for viewing.

5 DYNAMICAL HISTORY OF THE LOCAL GROUP IN ΛCDM – II. INCLUDING EXTERNAL PERTURBERS IN 3D (Banik & Zhao 2017)

To better understand if ΛCDM really faces a problem with the dynamics of LG galaxies, I followed up the work of Section 4 using a 3D model kindly lent by P. J. E. Peebles, who had previously used it to understand LG dynamics (Peebles & Tully 2013). The algorithm works by applying the numerical action method to solve the governing equations (Phelps et al. 2013). They tested the method by applying it to results of cosmological simulations, recovering the galaxy masses fairly well (see their figure 2).

5.1 Method

The 3D model lent by Peebles already included the LG galaxies and surrounding perturbers listed in table 1 of Peebles et al. (2011). This brightness-based catalogue misses the vast majority of the galaxies analysed in Section 4, a major shortcoming because LG dwarfs $\sim 1-3$ Mpc away turned out to be crucial to its conclusions. Thus, if not already present in the 3D model, I added the galaxies analysed in that work to the catalogue. LG dwarfs have very low masses, allowing me to add them as test particles satisfying the equation of motion

$$\ddot{\boldsymbol{r}} = {H_0}^2 \Omega_{\Lambda,0} \boldsymbol{r} - \sum_{\substack{\text{j = Distant}\\ \text{massive}\\ \text{particles}}} \frac{G M_j \left(\boldsymbol{r}-\boldsymbol{r}_j\right)}{|\boldsymbol{r}-\boldsymbol{r}_j|^3} - \sum_{\substack{\text{j = Nearby}\\ \text{massive}\\ \text{particles}}} \frac{G M_j \left(\boldsymbol{r}-\boldsymbol{r}_j\right)\left({r_c}^2+{r_{S,j}}^2\right)}{\left(|\boldsymbol{r}-\boldsymbol{r}_j|^2+{r_c}^2\right){r_{S,j}}^3} \tag{44}$$

When determining the force between any pair of massive galaxies, the same equation applied but I used the value of r_S corresponding to the galaxy with the larger r_S. The massive galaxies in this analysis are given in Table 4. The distances and HRVs shown are best-fitting values considering all observational constraints within their uncertainties (Section 5.2).

The gravitational field near massive particles is handled slightly differently than in my 2D model (Section 4). For any given test particle A, an explicit distinction is now drawn between massive particles whose r_S is below the distance to A and masses for which this is not the case, forces from which are handled using a pure inverse square law. Forces from nearby masses at first rise linearly with separation before falling as $g \propto \frac{1}{r}$, recovering the observed flat rotation curves of galaxies. The transition occurs near $r_c = 10$ kpc.

For the MW and M31, r_S is defined in the same way as in Section 4, though with the added assumption that the LSR speed is the same as v_f for the MW. Its value is allowed to float, with a prior assumption of 240 ± 10 km/s. The same value is used for M31. For other massive galaxies, I assume $r_S = 100$ kpc to avoid an adjustment each time their masses are altered.

Galaxy	Distance, Mpc	HRV, km/s	Mass, $10^{12} M_\odot$
Milky Way	0.008	−11.10	1.8302
Andromeda (Messier 31)	0.707	−309.18	2.0567
Centaurus A	3.736	504.52	5.8831
Messier 101	7.391	439.62	9.3108
Messier 94	4.366	324.31	8.8144
Sculptor	4.095	246.97	6.9296
NGC 6946	5.859	107.38	4.6142
Messier 81	3.625	73.48	4.0625
Maffei	3.988	−28.75	3.4924
IC 342	3.350	−12.98	1.2994
Triangulum (Messier 33)	0.948	−192.72	0.2214
Large Magellanic Cloud	0.065	235.97	0.2007
NGC 55	2.035	163.16	0.1323
NGC 300	1.963	158.70	0.1073
IC 10	0.781	−338.02	0.0437
NGC 185	0.706	−213.37	0.0129
IC 5152	1.878	138.56	0.0094
NGC 147	0.679	−201.04	0.0064
NGC 6822	0.510	−69.93	0.0059

Table 4. Data on the massive galaxies in my 3D model using a similar catalogue to Peebles et al. (2011, table 1). Distances and masses are allowed to vary to best match observations, though their prior distributions are not uniform (see text). The masses derived in my model correspond to the total halo mass of each system, some of which is located beyond its virial radius (Peñarrubia & Fattahi 2017). The top section of this table contains galaxies which are also directly included as massive extended objects in the 2D model (Section 4). The remaining galaxies are sorted in descending order of simulated mass. For clarity, I abbreviated the names of galaxies from the New General Catalogue (NGC) and Index Catalogue (IC).

Equation 44 is slightly different to that used in my 2D analysis of the LG (Equation 36). The matter portion of the cosmological acceleration $\ddot{a}$ is handled in a different way, though the dark energy is handled similarly. Previously, the Universe was treated as homogeneous except for a few massive particles. The cosmic expansion $\boldsymbol{r} \propto a$ could then be recovered at long range because $\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r}$. Here, I treated the Universe as empty except for the massive particles that I explicitly include. As the Universe is homogeneous on large scales, an accurate understanding of all the matter interior to a sufficiently distant test particle also leads to its separation from us changing with time as $\boldsymbol{r} \propto a$.

To see if this applies to my model, I determined how much mass was in the simulation out to the distance of M101, the most distant galaxy in my sample. The result of $4.9 \times 10^{13} M_\odot$ corresponds to a sphere of radius 7.01 Mpc filled with matter at a density equal to the present cosmic mean value. This is similar to the observed distance of M101 (Shappee & Stanek 2011), suggesting that the massive galaxies in my model mimic a smooth distribution on large scales with the correct density. A similar conclusion would be reached if I only considered the galaxies in my catalogue that lie within 3 Mpc of the MW. Thus, the matter distribution used should be accurate enough out to a sufficiently large distance that it enables the construction of an accurate dynamical model for motions within the LG.

The heart of this model involves solving the equations of motion by adjusting a trial trajectory towards the true one. An incorrect trajectory will have a mismatch between the

acceleration along it and that expected due to the gravity of other particles. Thus, at each timestep, the positions of all the particles are adjusted to try and equalise the gravitational field acting on each one with the acceleration $\ddot{\boldsymbol{r}}$ it experiences along its trajectory. This is done assuming both respond linearly to a position adjustment, although only the latter does. Thus, a solution can only be obtained after several iterations, each of which is reliant on a matrix inversion to handle the highly inter-connected nature of the problem. Some shortcuts are taken for test particles because their positions do not affect forces felt by other particles.

This method of solution is second-order accurate because of the standard finite differencing scheme used to obtain accelerations from a series of discrete positions valid at known times. Due to the large number of particle pairs, an adaptive timestep is impractical. Instead, I adapt the temporal resolution to the problem in a fixed way based on physical considerations. Each timestep corresponds to an equal increment in the cosmic scale-factor a, with 500 steps used between when $a = 0.1$ and the present time ($a \equiv 1$).

To check if I had adequate resolution, I solved the problem by forward Runge-Kutta integration instead, using 5000 timesteps equally spaced in a. The maximum error in the present position was 0.23 kpc while that in the velocity was 0.84 km/s. Both errors are very small, suggesting that there was enough resolution. Some other checks were also done to verify the numerical accuracy of the solution (Peebles et al. 2011, section 2.4).

5.2 Determining χ^2 and finding the best model

Like my axisymmetric model (Section 4), the 3D model accurately matches the observed sky positions of target galaxies. However, this is achieved rather differently. Instead of integrating the equations of motion forwards in time and using the Newton-Raphson method to very precisely match present positions, the 3D model integrates backwards in time starting from a position along the line of sight towards a target galaxy. I no longer require agreement between simulated and observed heliocentric distances. Instead, I add a contribution to the total χ^2 of the model if there is a mismatch. Handling distance uncertainties in this way makes error budgets model-independent, allowing relative model likelihoods to be determined simply by comparing their χ^2.

The distance errors σ_d come from observations. For M31, I use a slightly closer and more uncertain estimate (770 ± 40 kpc, Ma et al. 2010). Galaxies outside the LG might be affected by objects beyond the region covered by the analysis. It can also be difficult to determine the mass ratios between galaxies in an extended group and thus the location of its barycentre. For these reasons, I allow a fairly large distance uncertainty σ_d for such objects.

$$\frac{\sigma_d}{d_{MW}} = \frac{1}{10} \quad \text{if } d_{MW} > 3.2 \text{ Mpc} \tag{45}$$

Mismatches between observed and simulated GRVs are handled similarly, based on a tolerance of 20 km/s rather than the actual HRV measurement uncertainty. This is because I do not expect the model to be much more accurate as a representation of ΛCDM considering the level of scatter about the Hubble flow in more detailed cosmological simulations of the paradigm (Aragon-Calvo et al. 2011). As the observational uncertainty σ_{v_h} is always much smaller than this, the effect of raising it to 20 km/s is similar to setting $\sigma_{extra} = 20$ km/s in Equation 43. Either method handles modelling uncertainties by preventing the analysis from placing undue statistical weight on a very precisely observed galaxy.

Some LG galaxies have proper motion measurements. I made use of such data for M31, M33, the Large Magellanic Cloud (LMC), IC 10 and Leo I by adding a penalty to χ^2 when simulated and observed values disagree. Observational proper motion error estimates are taken at face value.

Unlike in my 2D model, Equation 35 is no longer strictly enforced at the start of the simulation because this is difficult to achieve when integrating backwards. Instead, I penalise models which fail to enforce it.

$$\Delta\chi^2 = \frac{|\overbrace{\boldsymbol{v}_i - H_i \boldsymbol{r}_i}^{\boldsymbol{v}_{pec}(t=t_i)}|^2}{{\sigma_v}^2} \tag{46}$$

Based on present-day deviations from the Hubble flow (Figure 16), I assume that the typical peculiar velocity $\boldsymbol{v}_{pec}$ was $\sigma_v = 50$ km/s when $a = 0.1$. This is a 1D measure which underestimates typical values of v_{pec} today. However, the nearly homogeneous state of the Universe at recombination (Planck Collaboration XIII 2016) implies that v_{pec} was typically smaller than today when my simulations started.

I do not fix the masses of any simulated galaxies which are treated as massive (Table 4). The prior used prefers a particular mass based on assuming M/L is 50× the Solar value $\left(\frac{M}{L_K}\right)_\odot$ in the near-infrared K-band (Tully et al. 2013). Observational estimates of the luminosity L_K in this band are based on a particular distance to each target. If the simulated distance is lower, then the model implies that the target is likely closer to us and thus intrinsically fainter for the same apparent magnitude. This makes it likely to be less massive. Accounting for this, I define a preferred mass estimate

$$M_c \equiv 50 L_K \left(\frac{M}{L_K}\right)_\odot \left(\frac{d_{model}}{d_{obs}}\right)^2 \tag{47}$$

Using a different mass M incurs a χ^2 penalty of

$$\Delta\chi^2 = \left[\frac{Ln\left(\frac{M}{M_c}\right)}{Ln\ 1.5}\right]^2 \tag{48}$$

For the MW and M31, there is no a priori preference towards any particular mass for either galaxy. However, a particular *ratio* between their masses is preferred. This is the ratio of their M_c values.

$$\Delta\chi^2 = \left(\frac{Ln\ \frac{M_{MW}}{M_{M31}} - Ln\ \frac{M_{MW,c}}{M_{M31,c}}}{Ln\ 1.25}\right)^2 \tag{49}$$

The model now has too many parameters to permit a grid search through them. Thus, I focus on results from the best-fitting 3D model, which I obtain by minimising χ^2 using a downhill-seeking walk through parameter space (Peebles et al. 2011, section 2.2). Each parameter A is varied by a small amount ΔA in an attempt to reduce χ^2. If this does not happen, then the algorithm restores the previous solution and reverses ΔA while also reducing its magnitude.

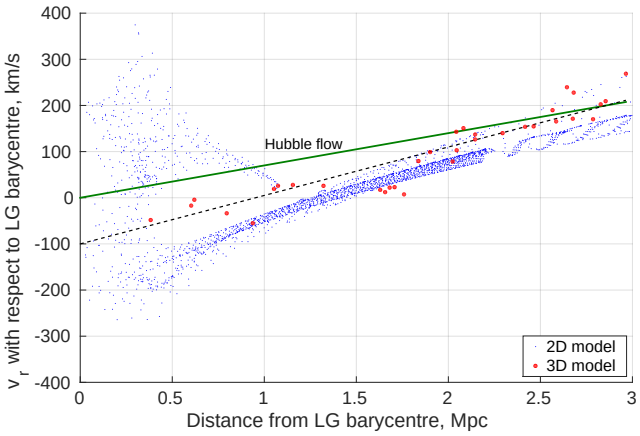


Figure 16. Radial velocities of test particles with respect to the LG barycentre are shown in blue for my 2D model with parameters matched to my best-fitting 3D model (Table 6), results of which are shown as large red dots. The solid green line is the Hubble flow relation for $H_0 = 70$ km/s/Mpc, the adopted value. The dashed black line has a gradient 1.5× larger. Due to the effect of gravity, it provides a better fit to the 3D model within the LG (Section 5.3).

Galaxy	ΔGRV (km/s)	Distance from LG barycentre (Mpc)
HIZSS 3	123.2 ± 10.6	1.76 ± 0.11
NGC 3109	110.7 ± 7.3	1.63 ± 0.05
Sextans A	95.1 ± 7.2	1.66 ± 0.02
Sextans B	75.4 ± 5.4	1.71 ± 0.05
Antlia	61.6 ± 8.3	1.68 ± 0.06
UGC 4879	-31.1 ± 5.5	1.32 ± 0.02
KKR 3	-33.6 ± 10.9	2.30 ± 0.12
GR 8	-40.0 ± 10.5	2.42 ± 0.12
NGC 55	-42.0 ± 10.4	2.08 ± 0.11
NGC 4163	-130.6 ± 7.7	2.96 ± 0.04

Table 5. ΔGRVs with respect to my 3D model for the 5 LG galaxies with the most positive and negative ΔGRVs (excluding NGC 404 and Leo P due to large distance uncertainties). Errors are estimated using Equation 43. The LG barycentre is put almost exactly at the MW-M31 mid-point (Table 6). Errors in the distance from there are obtained from those on heliocentric distances in the usual way. Notice how the tails of the ΔGRV distribution are rather asymmetric, as shown in Figure 18.

This is necessary because the increase in χ^2 is often due to overshooting the minimum.

$$\Delta A \;\rightarrow\; -\frac{1}{2}\Delta A \quad \left(\chi^2 \text{ increased}\right) \tag{50}$$

When a parameter adjustment reduces χ^2, I accelerate the convergence by setting

$$\Delta A \;\rightarrow\; \frac{5}{4}\Delta A \qquad \left(\chi^2 \text{ decreased}\right) \tag{51}$$

To avoid the parameter adjustments being too large or too small, a cap is imposed on $|\Delta A|$ such that

$$\left|\frac{\Delta A}{A}\right| \;<\; 10^{-1} \tag{52}$$

I assume the process has converged once $\left|\frac{\Delta A}{A}\right| < 10^{-5}$.

5.3 Results and discussion

A comparison between my best-fitting 2D and 3D models is complicated somewhat by the latter having many more degrees of freedom. In particular, it is not required to match the observed distances of LG galaxies. This allows it to place a galaxy further away than observed, increasing its predicted GRV and better explaining a very high observed GRV. I handle this by applying a correction to the predicted GRV of each galaxy based on how its simulated distance differs from the observed value. Thus, I set

$$GRV_{model} \;\rightarrow\; GRV_{model} + (d_{obs} - d_{model})\,\alpha H_0 \tag{53}$$

$$\alpha \;\equiv\; \frac{1}{H_0}\frac{dv_r}{dr} \tag{54}$$

I use $\alpha = 1.5$ because this provides a reasonable description of how radial velocities v_r depend on distances within the LG (Figure 16). At long range, $\alpha = 1$ but within the LG, gravity from the MW and M31 becomes important. Thus, an object further from them has been decelerated less by their gravity. Consequently, its radial velocity will be higher than for the more nearby object by a greater amount compared with a homogeneously expanding Universe. Neglecting projection effects (which become small a few Mpc from the LG), it is clear that α should slightly exceed 1.

In Section 4, I added an extra dispersion term σ_{extra} to Equation 43 and then marginalised σ_{extra} over other variables to obtain its probability distribution. The most likely value (using the optimal LMC mass) was 40.43 km/s. Using the same target galaxies, the rms dispersion in ΔGRV with respect to the best-fitting 2D model is 40.65 km/s, almost exactly the same. This suggests that the two statistics are very similar, even though the former uses integration over model parameter space while the latter is based on just one model. Thus, the rms ΔGRV of the best-fitting model should provide a very good guide to the results of a more thorough statistical analysis attempting to pin down how inaccurate each model is as a representation of the data. To draw conclusions about the validity of ΛCDM, this would then have to be compared with how accurately the models can be expected to represent ΛCDM.

After obtaining corrected GRV predictions for my 3D model using Equation 53, I subtracted them from observed GRVs (Equation 42) to obtain a list of ΔGRVs. The rms of these ΔGRVs is then found for a range of plausible assumptions regarding α (Figure 17). For comparison, I also show the result of the same calculation for my best-fitting 2D model using the same target galaxies. This model requires an extremely precise match between their simulated and observed distances, making the result independent of α. Although it was technically difficult to operate the 3D model in this way, one can gain a conservative lower bound on the rms value of ΔGRV had this been done by setting $\alpha = 0$ in Equation 53. This corresponds to taking the GRV predictions of the 3D algorithm at face value, even though it has some flexibility with distances. Removing this flexibility can only worsen the agreement between predicted and observed GRVs.

With α irrelevant for the 2D model, its main uncertainty is whether Cetus and DDO 216 should be included in the analysis as they are very discrepant with this model. I

Parameter	Meaning & units	Best-fitting value in 2D	Best-fitting value in 3D
M	LG mass, $10^{12} M_\odot$	2.756	4.088
q_{MW}	$M_{MW} \div M$	0.356	0.497
q_{LMC}	$M_{LMC} \div M_{MW}$	0.157	0.099
$v_{c,\odot}$	LSR speed, km/s	239	223.0
$v_{f,M31}$	v_f of M31, km/s	225	240.3
d_{M31}	Distance to M31, kpc	783	707
$M_{\rm Cen\ A}$	Cen A mass, $10^{12} M_\odot$	4	5.883
$U_\odot$	Components of the	14.1	11.1
$V_\odot$	non-circular motion of	14.6	12.2
$W_\odot$	Sun in the MW, km/s	6.9	7.2
H_0	Hubble constant	67.3	70
$\Omega_{m,0}$	Present matter density in the Universe $\div \frac{3{H_0}^2}{8\pi G}$	0.315	0.27

Table 6. The parameters of my best-fitting axisymmetric (2D) and 3D models. q_{LMC} is the LMC mass as a fraction of the MW mass, which I take to include the LMC. The top section of this table contains the parameters I varied using a grid search in my 2D model (Section 4) or using gradient descent in 3D (Section 5). The central section contains the parameters associated with the non-circular motion of the Sun in the Milky Way, which I obtain from Francis & Anderson (2014) for the 2D model and from Schönrich (2012) for the 3D model. This section also contains two parameters related to M31. In the 2D model, its distance estimate is from McConnachie (2012) while the 3D model uses a prior of 770 ± 40 kpc (Ma et al. 2010). Its rotation curve flatlines at a level $v_{f,M31}$ which is fixed in the 2D model but has a prior of 240 ± 10 km/s in the 3D model (Carignan et al. 2006). This model assumes $v_f = v_{c,\odot}$ for the MW whereas the 2D model fixes the former at 180 km/s (Kafle et al. 2012) and uses a prior on the latter of 239 ± 5 km/s (McMillan 2011). I adopt a flat dark energy-dominated cosmology whose parameters are fixed at the values given in the bottom section, with the 2D results based on those of Planck Collaboration XIII (2016) while the 3D results are based on Komatsu et al. (2011). Both models start when the cosmic scale-factor $a = 0.1$.

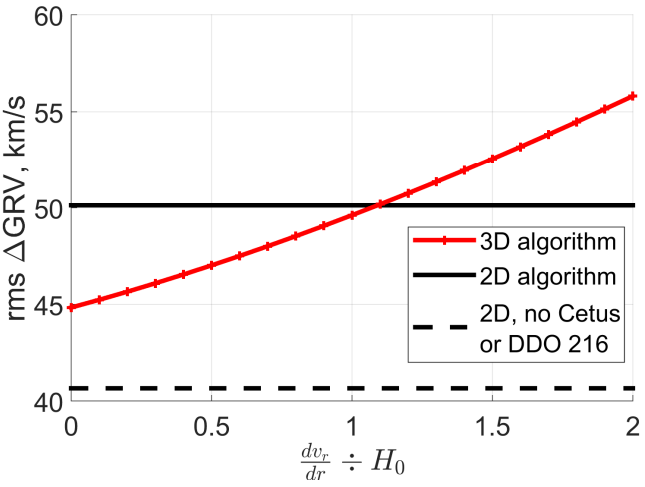


Figure 17. The root mean square value of ΔGRV for my best-fitting 2D (black) and 3D (red) models as a function of α, which governs how 3D model predictions are adjusted to put them on an equal footing with my 2D model (Equation 53). The adjustment is unnecessary for the latter. This model is likely unreliable close to M31 as it lacks M33, making its predictions for Cetus and DDO 216 unreliable. Results including these galaxies (solid black) and without them (dashed black) are shown. In the 3D model, removing them increases the results by only $\sim$0.7 km/s, thus leaving them almost unchanged. This model treats the LG as empty apart from a few point masses. Using a similar assumption in my 2D models would reduce the rms value of ΔGRV by $\sim$6 km/s (not shown).

suggest that they should not be included as they are quite close to M31. Unlike in the 2D case, excluding them from the 3D analysis hardly affects its rms value of ΔGRV (this rises $\sim$0.7 km/s), which then greatly exceeds the 2D result for the same sample. Even if these galaxies are included in both models, any value of $\alpha > 1.1$ implies that the rms ΔGRV is larger in the 3D analysis. Thus, modelling the LG in 3D does not alleviate the discrepancy with ΛCDM first highlighted using my 2D model (Section 4). In fact, the discrepancy is slightly worse in the 3D case.

To better characterise this discrepancy, the residuals between model predictions and observations (ΔGRVs) are shown as a histogram in Figure 17. These results show a similar pattern to my 2D results (Figure 14) in that the ΔGRV distribution for galaxies with $\Delta GRV < 0$ (shown in blue) can broadly be understood using a Gaussian of width 25 km/s, a reasonable estimate of the modelling uncertainty (e.g. due to neglecting interactions between LG dwarfs). The only exception is NGC 4163, a very distant LG galaxy whose observed GRV is much lower than nearby galaxies at similar heliocentric distances in the Canes Venatici I cloud (Makarov et al. 2013, table 2). NGC 4163 may have been flung towards the LG by a close interaction outside it that was not captured by my model.

Unlike galaxies with $\Delta GRV < 0$, those with $\Delta GRV > 0$ (shown in red) have a ΔGRV distribution completely different from a 25 km/s Gaussian. This discrepancy is not due to just one galaxy. To emphasise this, I use Table 5 to list the five galaxies with the highest and lowest (most negative) ΔGRVs compared to my best-fitting 3D model. The inferred LG parameters in this model are given in Table 6, which compares the results to those of my best-fitting 2D model. Neither model matches LG observations particularly well.

As discussed in Section 4, the kinematics of the HVGs might be due to enhanced gravitational slingshot interactions with the MW and M31 around the time of their flyby. If this is correct, ΔGRV should be larger for galaxies further from the LG. A trend of $u \sim 50$ km/s/Mpc is apparent in Figure 19 (dashed line), suggesting a MW-M31 flyby $\sim (H_0 + u)^{-1} = 8$ Gyr ago. This is roughly when the MW-M31 flyby is expected to have taken place in MOND (Zhao et al. 2013) and when the Galactic thick disk formed (Quillen & Garnett 2001). In Section 6, I consider other patterns that would be expected in this scenario.

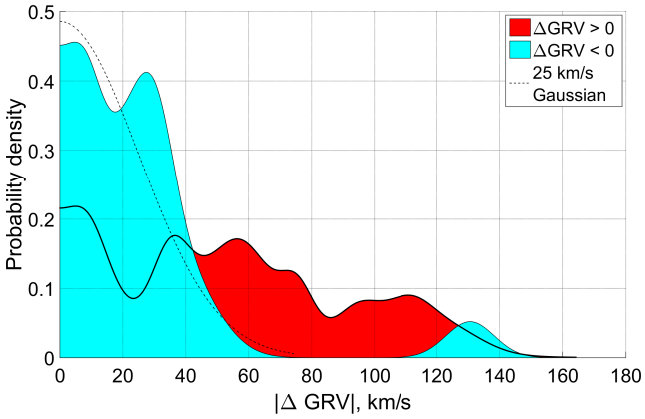


Figure 18. Histogram of ΔGRVs with respect to my 3D model, shown separately according to the sign of ΔGRV. The area of each square corresponds to 2 galaxies. A similar pattern emerges to my 2D model of the LG (Figure 14), with the blue bump near 130 km/s caused by NGC 4163. Otherwise, the galaxies with $\Delta GRV < 0$ (solid blue) are well described by a 25 km/s Gaussian (dashed line). This is not the case for galaxies with $\Delta GRV > 0$ (solid red).

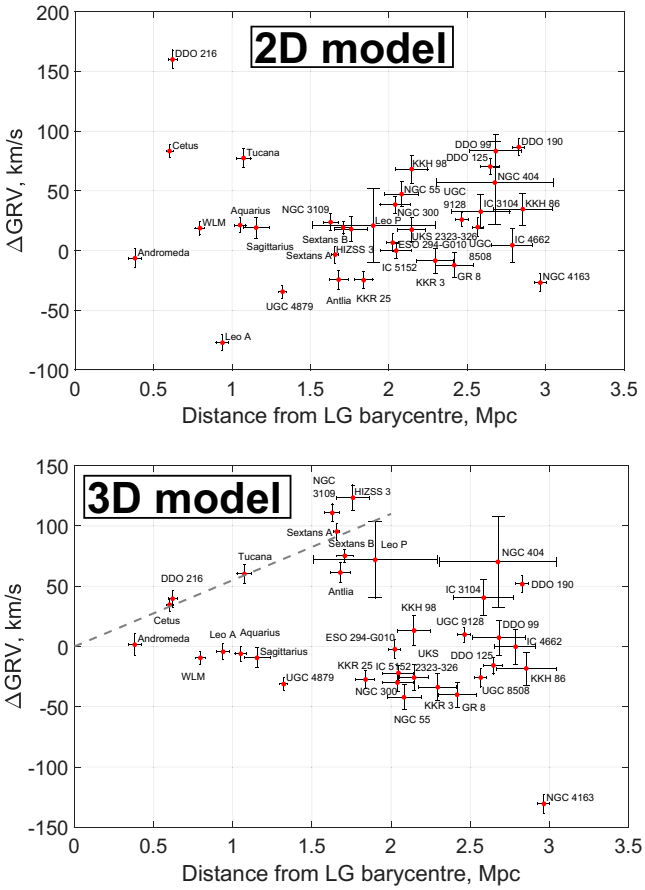


Figure 19. The ΔGRV of each target galaxy is shown against its distance from the LG barycentre. Parameters of the models used are given in Table 6, with the best-fitting ones used for the relevant number of dimensions in each model. Errors shown tend to be anti-correlated because a larger distance to a target increases its predicted GRV, reducing its ΔGRV.

6 A PLANE OF HIGH-VELOCITY GALAXIES ACROSS THE LOCAL GROUP (Banik & Zhao 2018b)

Section 5 confirmed my earlier result (Section 4) that the LG has several galaxies with unusually high radial velocities compared to ΛCDM expectations. In this section, I revisit that analysis with a more thorough search for model parameters and dwarf trajectories that agree better with observations (Section 6.1). As this still leaves several HVGs, I try and understand them using a MOND model of the LG (Section 6.2). A past MW-M31 flyby could yield HVGs at roughly the observed radial velocities, but it would tend to concentrate the HVGs within a particular plane. I consider how to test this statistically in Section 6.3 and conduct this test in Section 6.3.2, where I also consider other properties expected of the HVGs in this scenario (Table 9).

6.1 Revisiting the Local Group in ΛCDM

To better identify which galaxies may have been flung out by a fast-moving MW/M31, I refine the LG timing argument analysis of Section 4. The input catalogue is updated, with the main changes being a more accurate distance to NGC 404 (Dalcanton et al. 2009) and Leo P (McQuinn et al. 2015). For NGC 4163, I use a less accurate distance of 2.95 ± 0.07 Mpc to bracket the range between different Hubble-based measurements (Dalcanton et al. 2009; Jacobs et al. 2009).

The model is improved by relaxing the assumption that the flatline level of the MW rotation curve $v_{f,MW}$ is equal to its amplitude $v_{c,\odot}$ at the position of the Sun. Instead, I let $v_{f,MW}$ vary with a prior of 205 ± 10 km/s (McGaugh 2016a) while $v_{c,\odot}$ is fixed at 232.8 km/s (McMillan 2017). The time resolution is improved 10× so that the history of the Universe since redshift 9 ($a = 0.1$) is covered with 5000 steps, allowing for a much better handling of close encounters.

I also improved the χ^2 minimisation procedure, which is now done by applying gradient descent to all model parameters, as described in Section 6.3.1. To maximise the chance of matching observations, I ran a grid search through the trajectories of all the dwarf galaxies (treated as test particles). As explained in Section 5.1, trajectories were solved by relaxing an initial guess towards a solution that satisfies the equations of motion. The initial guess has the co-moving position varying linearly with a. Each dwarf's current $\boldsymbol{v}_{pec}$ was varied over a 3D grid of possibilities, giving the algorithm a much better chance of finding close encounters that might otherwise get missed if the initial trajectory went nowhere near the spacetime location of the encounter. Thus, the grid search is complementary to gradient descent, which can find the minimum more precisely but is more prone to finding a local minimum rather than the true global minimum.

As some improvements were indeed found in this way, I repeated the gradient descent stage and the grid search in an alternating manner until the algorithm converged (i.e. the grid search did not further reduce χ^2). This process took a few days and yielded reliable trajectories for all simulated galaxies − their present positions and velocities were almost perfectly recovered compared to a forward integration using

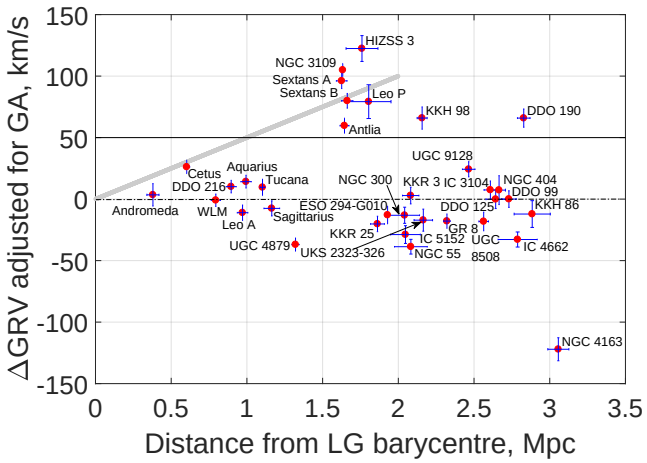


Figure 20. The deviation ΔGRV of each target galaxy from my best-fitting ΛCDM model, shown against its distance from the LG barycentre. An adjustment is applied to account for tides raised by the Great Attractor (Banik & Zhao 2017, equation 30). If the model worked perfectly, then all galaxies would have $\Delta GRV \equiv 0$ as model predictions are subtracted. Given likely model uncertainties of $\sim$25 km/s (Aragon-Calvo et al. 2011), ΛCDM would thus find it difficult to explain galaxies with $\Delta GRV > 50$ km/s. In the MOND scenario of a past MW-M31 flyby, the HVGs should broadly follow a trend of 50 km/s/Mpc (diagonal grey line) and reach distances up to $\sim$2 Mpc (Figure 22).

the fourth-order Runge-Kutta method with 10× finer resolution (maximum errors of 9 pc and 16 m/s, respectively).

Using equation 30 from Banik & Zhao (2017), I adjusted the predictions of this best-fitting model for the effect of tides raised on the LG by the Great Attractor (GA). This only slightly affects the results, which are shown in Figure 20. Compared to Section 4, the main difference is that Tucana is now consistent with ΛCDM expectations. Given modelling uncertainties, this applies to any galaxy with $|\Delta GRV| < 50$ km/s. However, several galaxies still have a larger ΔGRV.

In the MOND flyby scenario, the HVGs passed close to the spacetime location of the MW-M31 flyby. Thus, the HVGs should follow a $\Delta GRV \propto d$ relation of the sort marked on Figure 20. DDO 190 does not fit neatly into this picture. However, given 34 target galaxies, it would not be particularly unusual to find one with a ΔGRV as large as the 66 ± 7 km/s value for DDO 190 if the model uncertainty is taken to be 25 km/s (probability $P = 12\%$). A second such instance would be unexpected ($P = 0.8\%$). Bearing this in mind, I obtain the HVG sample listed in Table 7. This allows me to quantify the statistical properties of the HVGs and thereby better understand how their unusual kinematics arose (Section 6.3).

Galaxies included in the plane fit	Distance from MW-M31 mid-point, Mpc	ΔGRV, km/s
Milky Way	0.382 ± 0.04	NA
Andromeda	0.382 ± 0.04	3.5 ± 9.1
Sextans A	1.624 ± 0.036	96.1 ± 6.3
Sextans B	1.661 ± 0.037	79.9 ± 6.0
NGC 3109	1.631 ± 0.014	105.0 ± 5.3
Antlia	1.642 ± 0.030	59.7 ± 6.1
Leo P	1.80 ± 0.15	79 ± 14
KKH 98	2.160 ± 0.033	65.5 ± 9.1

Table 7. Galaxies considered HVGs based on Figure 20. The MW and M31 are shown here for reference.

6.2 Simulating the Local Group in MOND

6.2.1 Governing equations

As in ΛCDM models of the LG, it is first necessary to obtain the MW-M31 trajectory. This is done by advancing them according to their mutual gravity supplemented by the cosmological acceleration term (e.g. Equation 36).

$$\ddot{\boldsymbol{r}}_{rel} = \boldsymbol{g}_{M31} - \boldsymbol{g}_{MW} + \frac{\ddot{a}}{a}\boldsymbol{r}_{rel} \quad \text{where} \tag{55}$$

$$\boldsymbol{r}_{rel} \equiv \boldsymbol{r}_{M31} - \boldsymbol{r}_{MW} \tag{56}$$

$\boldsymbol{r}_i$ is the position vector of galaxy i (MW or M31), at whose location the gravitational field (excluding self-gravity) is $\boldsymbol{g}_i$. All position vectors are with respect to the LG barycentre, which I take to be 0.3 of the way from M31 towards the MW. This is based on the MW rotation curve asymptotically reaching a flat level of $\sim$180 km/s (Kafle et al. 2012) while the equivalent value for M31 is $\sim$225 km/s (Carignan et al. 2006). In the context of MOND, this suggests that the mass of M31 is $\left(\frac{225}{180}\right)^4 \approx 2.3\times$ that of the MW (Equation 10).

To find the gravitational field $\boldsymbol{g}$ due to the MW and M31 at some position $\boldsymbol{r}$, I treat them as point masses and use the quasilinear formulation of MOND (QUMOND, Milgrom 2010). In spherically symmetric situations, this yields identical forces to the more traditional aquadratic Lagrangian formulation (AQUAL, Bekenstein & Milgrom 1984). Even in less symmetric systems, the forces differ by only a few percent (e.g. Banik & Zhao 2015b; Candlish et al. 2016). As discussed in Section 1.5, I use the 'simple' interpolating function between the Newtonian and deep-MOND regimes.

$$\boldsymbol{g}_N \equiv -\sum_{i=MW,M31} \frac{GM_i\left(\boldsymbol{r}-\boldsymbol{r}_i\right)}{\left|\boldsymbol{r}-\boldsymbol{r}_i\right|^3} \tag{57}$$

$$\nabla\cdot\boldsymbol{g} \equiv \nabla\cdot\left[\nu\left(\frac{|\boldsymbol{g}_N|}{a_0}\right)\boldsymbol{g}_N\right] \quad \text{where} \tag{58}$$

$$\nu\left(x\right) = \frac{1}{2} + \sqrt{\frac{1}{4}+\frac{1}{x}}$$

For an isolated point mass, $\boldsymbol{g}$ is given by Equation 8. In the more complicated axisymmetric situation relevant to this problem, I use direct summation to obtain $\boldsymbol{g}$ from its divergence.

$$\boldsymbol{g}\left(\boldsymbol{r}\right) = \int \nabla\cdot\boldsymbol{g}\left(\boldsymbol{r}'\right)\frac{\left(\boldsymbol{r}-\boldsymbol{r}'\right)}{4\pi\left|\boldsymbol{r}-\boldsymbol{r}'\right|^3}\,d^3\boldsymbol{r}' \tag{59}$$

$\nabla\cdot\boldsymbol{g}$ is calculated out to almost 150 r_{rel}, beyond which it should be very nearly spherically symmetric. Due to the shell theorem, it is unnecessary to consider $\nabla\cdot\boldsymbol{g}$ (or 'phantom dark matter') at even larger radii. As $\boldsymbol{g}$ is only determined out to 66.5 r_{rel}, my results should be nearly free of edge effects. At larger distances, the MW and M31 are treated as a single point mass located at their barycentre, yielding $\boldsymbol{g} = \nu\boldsymbol{g}_N$ (Equation 8).

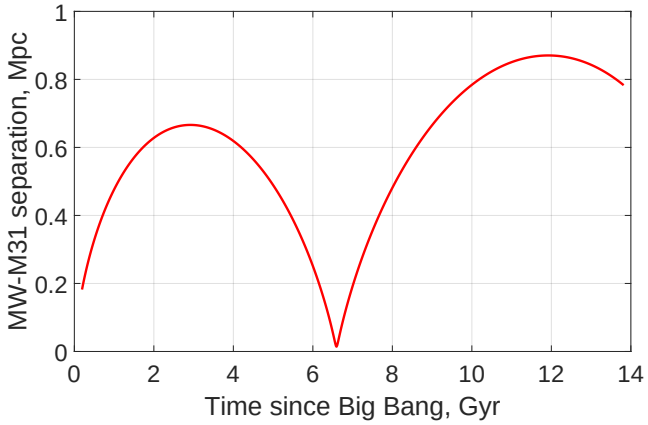


Figure 21. MW-M31 separation in my MOND simulation, showing a past close flyby 6.59 Gyr after the Big Bang at a closest approach distance of 14.17 kpc. At that time, their relative velocity was 716 km/s, of which 501 km/s was due to motion of the MW. The higher second apogalacticon is partly due to the effect of cosmology (Equation 55) and my assumption that the MW and M31 lose 5% of their mass around the time of their encounter. Some other reasons for this are explained in section 5.1.2 of Banik et al. (2018).

In general, the galaxies will not be on the Hubble flow at the start time of the simulations t_i, when the cosmic scale-factor $a_i = 0.05$ and $H \equiv H_i$. However, deviations from the Hubble flow are observed to be very small at early times (Planck Collaboration XXVII 2014). In order to satisfy this condition, I varied the total MW and M31 mass M using a Newton-Raphson root-finding algorithm to ensure that

$$\dot{\boldsymbol{r}}_{rel} = H_i \boldsymbol{r}_{rel} \quad \text{when } t = t_i \tag{60}$$

The MW and M31 are not on a purely radial orbit. Their mutual orbital angular momentum prevents them from converging onto the Hubble flow at very early times. This is unrealistic as any non-radial motion must have arisen due to tidal[1] torques well after the Big Bang. Thus, I take the MW-M31 orbit to be purely radial prior to their first turnaround at $t \approx 3$ Gyr. After this time, I assume their trajectory conserves angular momentum at its present value. This implies the MW-M31 angular momentum was gained near the time of their first turnaround, when their large separation would have strengthened tidal torques. At later times, the larger scale factor would weaken tidal torques, suggesting that these have a much smaller effect around the time of the second MW-M31 turnaround than the first.

The resulting MW-M31 trajectory is shown in Figure 21. With this information, $\boldsymbol{g}$ can be found everywhere within the LG at all times using Equation 59, assuming only these point masses are present in an otherwise homogeneous Universe.

Test particle trajectories can now be advanced using

$$\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r} + \boldsymbol{g} \tag{61}$$

$$\dot{\boldsymbol{r}} = H_i \boldsymbol{r} \quad \text{when } t = t_i \tag{62}$$

[1] affecting the MW and M31 differently

Because the MW and M31 must have accreted matter from some region prior to the start of the simulation, I exclude all test particles starting within a distance $r_{exc,i}$ of galaxy i. This distance is determined by requiring that the excluded volume has as much baryonic matter as galaxy i, taking the density of baryons to be the cosmic mean value. To obtain this, I assume baryons currently comprise a fraction $\Omega_{b,0} = 0.049$ of the cosmic critical density, which I found by taking $H_0 = 67.3$ km/s/Mpc (Planck Collaboration XIII 2016, table 4). The cosmic baryon density can be estimated using Big Bang nucleosynthesis – only a narrow range of values is consistent with the primordial abundances of light elements like deuterium (Cyburt et al. 2016).

$$\frac{4\pi}{3} {r_{exc,i}}^3 \times \overbrace{\frac{3{H_0}^2}{8\pi G}\Omega_{b,0} {a_i}^{-3}}^{\text{Baryon density at } t_i} \equiv M_i \quad (\text{for } r_{exc,i}) \tag{63}$$

To avoid the excluded regions overlapping, it is necessary that their sizes satisfy

$$r_{exc,MW} + r_{exc,M31} \leqslant r_{rel} \quad \text{when } t = t_i \tag{64}$$

This inequality applies because r_{exc} is 77.7 kpc for the MW and 102.5 kpc for M31, leading to a total of 180.2 kpc – interestingly, this is just smaller than $r_{rel}(t_i) = 182.1$ kpc, suggesting that the two galaxies accreted matter from regions which just touched. This remains the case with a slightly different start time as $r_{exc,i} \propto a_i$, similarly to $r_{rel}(t_i)$ – at such early times, both galaxies would follow the Hubble flow rather closely. However, this coincidence does not occur in ΛCDM, a model in which the excluded regions would very likely overlap (Banik & Zhao 2016, section 2.2.1).

6.2.2 Results

Figure 22 shows the distances and radial velocities of test particles with respect to the LG barycentre, colour-coded according to the orientation of their orbital plane relative to that of the MW-M31 orbit. This is quantified based on each particle's specific angular momentum $\boldsymbol{h}$, whose direction can readily be compared with the MW-M31 orbital pole $\widehat{\boldsymbol{h}}_{LG}$.

$$\boldsymbol{h} \equiv \boldsymbol{r} \times \dot{\boldsymbol{r}} \tag{65}$$

$$\cos\psi \equiv \widehat{\boldsymbol{h}} \cdot \widehat{\boldsymbol{h}}_{LG} \tag{66}$$

For this section, the important feature is the upper branch of the Hubble diagram. Its upward slope arises because these particles must have passed close to the spacetime location of the MW-M31 encounter and gained a substantial amount of kinetic energy in what was essentially a 3-body interaction. Thus, for such particles to be further away from the LG now, they must have a larger outwards velocity.

ΛCDM also allows slingshot encounters with the MW and M31, but their fairly slow motion only allows them to fling galaxies out to ~ 1 Mpc from the LG. At this point, the upper branch of the Hubble diagram simply stops (Figure 12). Even in more detailed cosmological simulations of ΛCDM that include encounters with satellites of MW and M31 analogues, dwarf galaxies do not get flung out beyond this distance (Sales et al. 2007, figures 3 and 6). For MOND, the corresponding limit is ~ 2.5 Mpc due to the MW-M31 flyby, which therefore makes a dramatic difference to the Hubble diagram at distances of $\sim 1-2$ Mpc (Figure 22).

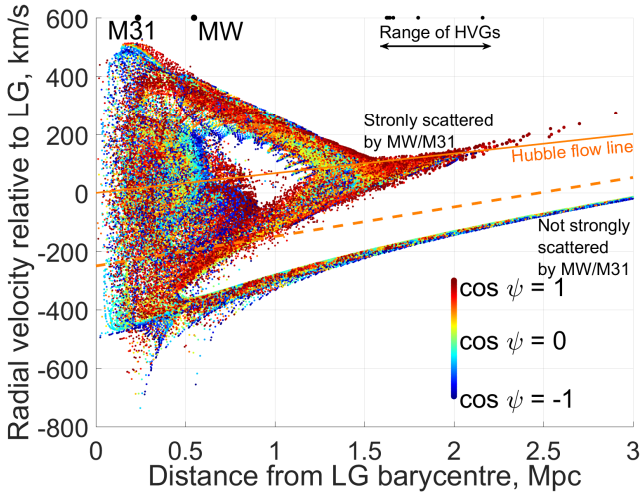


Figure 22. Hubble diagram of the test particles in my simulation coloured by their value of $\cos\psi$, which parametrises how well their orbital angular momenta align with that of the MW-M31 orbit (Equation 66). The Hubble flow line is shown in solid orange. I also show a 1.5× steeper line (dashed orange) that is used to select analogues of HVGs in Figure 23. Particles below this line have generally never interacted closely with the MW or M31, unlike particles above the line. The black dots along the top edge of the figure indicate distances to the MW, M31 and the HVGs (Table 7).

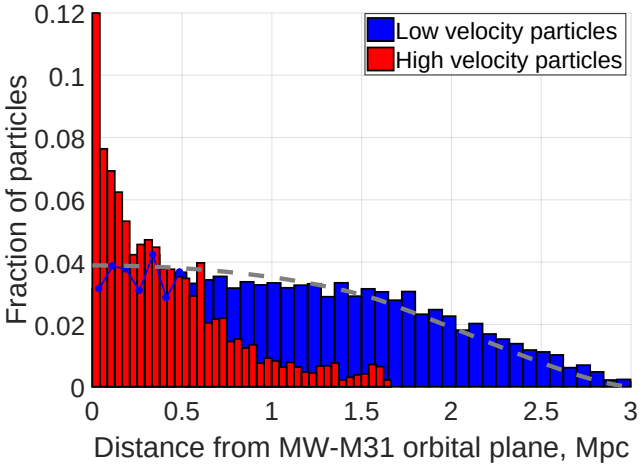


Figure 23. Histogram showing how far simulated particles are from the MW-M31 orbital plane. This only shows particles currently $1.6-3$ Mpc from the LG, sorted according to whether they are in the high-velocity branch of the Hubble diagram (above dashed orange line in Figure 22). If they are, I show them as red. The remaining particles (shown in blue) are well described by an isotropic distribution (dashed grey line).

In this distance range, my simulation yields a bimodal distribution of radial velocities, with the HVGs corresponding to particles in the upper branch. A pattern of this sort is apparent in the kinematics of the observed LG (Figure 20).

Gravitational slingshot interactions with the MW or M31 would be most efficient for particles flung out roughly parallel to the motion of the perturber. Considering that the MW-M31 flyby occurred a fixed time in the past, these particles should currently be furthest away from the LG. Thus, it is not very surprising that the spatial distribution of such distant HVG analogues is indeed highly flattened with respect to the MW-M31 orbital plane (Figure 23).

6.3 Statistical analysis

6.3.1 Finding the best-fitting plane

To quantify whether a set of galaxies is distributed anisotropically, it is necessary to define a measure of anisotropy and determine how unusual its value is. The statistic I used was z_{rms}, the rms of the minimum distances between the galaxies I consider and the best-fitting plane through them (i.e. the one that minimises z_{rms}). With respect to a plane having normal $\widehat{\boldsymbol{n}}$ and containing the vector $\boldsymbol{r}_0$, the vertical dispersion is

$$z_{rms}{}^2 = \frac{1}{N}\sum_{i=1}^{N}\left[(\boldsymbol{r}_i-\boldsymbol{r}_0)\cdot\widehat{\boldsymbol{n}}\right]^2 \quad (67)$$

$$= \widehat{\boldsymbol{n}}\cdot(\mathbf{I}\widehat{\boldsymbol{n}}) \quad \text{where} \quad (68)$$

$$\mathbf{I}_{jk} \equiv \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{r}_i-\boldsymbol{r}_0)_j(\boldsymbol{r}_i-\boldsymbol{r}_0)_k \quad (69)$$

Quantity	Full sample	Without Antlia
Galaxies in plane	8	7
Normal to plane of high ΔGRV galaxies	$\begin{bmatrix}204.4^\circ\\-30.1^\circ\end{bmatrix}$	$\begin{bmatrix}206.6^\circ\\-31.8^\circ\end{bmatrix}$
rms plane width, kpc	101.1	101.9
Aspect ratio (Eq. 71)	0.0763	0.0750
MW offset from plane	224.7	195.4
M31 offset from plane	-0.6	-12.8
Angle of MW-M31 line with plane	16.2°	14.9°

Table 8. Information about the plane best fitting the galaxies listed in Table 7, with distances in kpc and plane normal direction in Galactic co-ordinates (latitude last). The last column shows how the results change if Antlia is removed from the sample as it could be a satellite of NGC 3109 (van den Bergh 1999).

The galaxies are at heliocentric positions $\boldsymbol{r}_i$. The minimum of z_{rms} is attained when $\boldsymbol{r}_0 = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{r}_i$, corresponding to the geometric centre of the N galaxies to which a plane is being fit. I find its best-fitting orientation $\widehat{\boldsymbol{n}}$ using a gradient descent method (e.g. Fletcher & Powell 1963). Issues of local minima are resolved by starting the gradient descent at whichever $\widehat{\boldsymbol{n}}$ yields the smallest z_{rms} in a low resolution grid of possible directions for $\widehat{\boldsymbol{n}}$. Once the angular step size is below 0.006°, further iterations are stopped and I assume the algorithm has converged to an acceptable precision.

Applied to the major LG galaxies along with the HVGs except HIZSS 3 (full list in Table 7), this reveals that they define a rather thin plane whose parameters are given in Table 8. This allows a comparison between the ΔGRV of

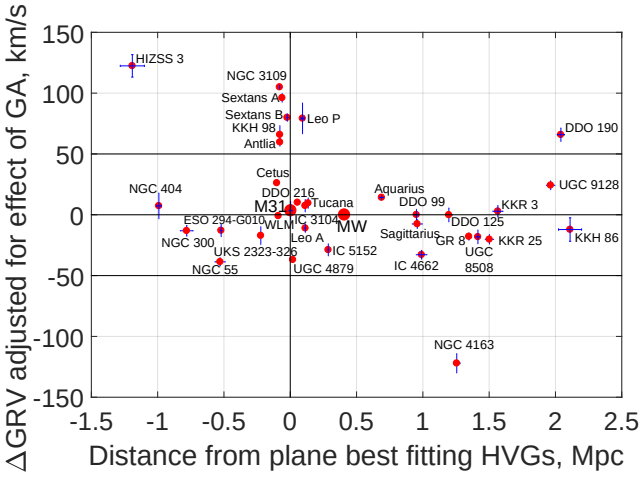


Figure 24. ΔGRVs of target galaxies are shown against their offsets from the best-fitting plane through the ones with the largest ΔGRVs except HIZSS 3 (parameters of this plane given in central column of Table 8). By definition, ΛCDM predicts $\Delta GRV = 0$ with an uncertainty of $\sim$25 km/s (Aragon-Calvo et al. 2011). Thus, it can't easily explain galaxies with $\Delta GRV > 50$ km/s (above upper horizontal line). Most of these galaxies lie very close to a plane (near vertical gridline), unlike the rest of the sample. The concept of a ΔGRV is meaningless for the MW, so I show this as 0.

each galaxy[1] and its minimum distance from this plane. The galaxies in the full sample have a wide range of positions relative to it, with a similar number on either side (Figure 24). However, the HVGs tend to lie very close to it. The only exception is HIZSS 3, justifying my decision not to consider it when defining the HVG plane. In any case, the observations for HIZSS 3 are rather insecure due to its very low Galactic latitude (0.09°, Massey et al. 2003). Some of the issues caused by this are discussed in section 6.3 of Banik & Zhao (2018b). Apart from HIZSS 3, my analysis has no target galaxies within 16° of the Galactic plane.

6.3.2 Monte Carlo trials

To see if the situation shown in Figure 24 is consistent with isotropy, I conduct a series of Monte Carlo (MC) trials in which I randomise the sky directions of the HVGs and recompute z_{rms}. Formally, isotropy implies that the Galactic longitude l has a uniform probability distribution while that of the Galactic latitude b is

$$P(b)\ db\ =\ \frac{1}{2}\cos b\ db \tag{70}$$

To mimic uncertainties in measured distances to LG galaxies, I randomly vary their heliocentric distances using Gaussian distributions of the corresponding widths. Although this happens very rarely, any negative distances are raised to 0.

To account for HIZSS 3 being excluded from the plane fit despite its high ΔGRV, I used the procedure described in Section 6.3.1 to find the best-fitting plane through every combination of all HVGs but one as well as the MW and M31. The combination yielding the lowest z_{rms} is considered the analogue of the observed HVG system less HIZSS 3 for that particular randomly generated mock catalogue. The enforced inclusion of the MW and M31 is necessary because it would not make sense for them to lie far from a plane supposedly corresponding to their mutual orbital plane.

To check whether the particular choice of statistic affected my results, I also performed calculations where I selected the combination yielding the lowest aspect ratio A rather than z_{rms}.

$$A\ \equiv\ \frac{z_{rms}}{\sqrt{{r_{rms}}^2-{z_{rms}}^2}}\quad\text{where} \tag{71}$$

$${r_{rms}}^2\ \equiv\ \frac{1}{N}\sum_{i=1}^{N}|\boldsymbol{r}_i-\boldsymbol{r}_0|^2\ =\ Trace\,(\mathbf{I}) \tag{72}$$

r_{rms} is the rms distance of the galaxies from their geometric centre $\boldsymbol{r}_0$. To get the rms extent of the system after projection into the best-fitting plane, I need to subtract z_{rms} in quadrature. Dividing z_{rms} by the result then gives a measure of the typical 'vertical' extent of galaxies out of this plane relative to their 'horizontal' extent within it. Identical probabilities would be obtained had I defined $A \equiv \frac{z_{rms}}{r_{rms}}$ instead, as long as it is defined in the same way for the actual HVGs and the mock sample in each MC trial. This is because A is a monotonic function of $\frac{z_{rms}}{r_{rms}}$ with either definition. Moreover, both definitions yield very similar A for a highly flattened structure.

In Table 9, I give the criteria which I used to determine whether the HVGs in each MC trial are distributed in an analogous way to observations. I choose these criteria so that they should be satisfied if the LG behaves similarly to my MOND simulation of it (Section 6.2). I consider one of the first two anisotropy-related criteria alongside both of the others. I used a toy model to crudely estimate the MW-M31 orbital pole required to explain the observed orientations of their satellite planes (Banik & Zhao 2018b, section 2.2). The MW-M31 orbital pole preferred by this toy model has only a small impact on my final results, which would be similar if the constraint it yields was not considered (Table 10).

If a past MW-M31 flyby is responsible for the unusual kinematics of the HVGs, then the plane they define should intersect the MW-M31 barycentre. I take this to be 0.3 of the way from M31 towards the MW for reasons discussed at the start of Section 6.2.1. This puts the MW-M31 barycentre 67 kpc from the best-fitting plane, a rather small offset from a plane with a radial extent of $\sim$1.3 Mpc.

6.4 Results

Applying the criteria defined in Table 9 to 20 million MC trials based on my nominal sample of HVGs (Table 7), I obtained the results shown in Table 10. The uncertainties are found by repeating the MC trial using 4 different seeds for the random number generator, with each seed used for 5×10^6 trials. The variance between the results gave an indication of the uncertainty in the final result, which is a simple mean. I also estimated the error using binomial statistics. My final error estimate was based on whichever method gave a higher uncertainty (this was usually based

[1] adjusted for the Great Attractor using equation 30 of Banik & Zhao (2017)

Criterion	Meaning
Plane thickness	There must be a plane of HVGs with rms thickness (Equation 67) below that observed
Aspect ratio	There must be a plane of HVGs with aspect ratio (Equation 71) below that observed
Barycentre offset	Barycentre of MW and M31 (assuming 30% of total mass in MW) closer to plane than observed situation
Direction	Normal to HVG plane closer than observed to expectation of toy model

Table 9. Criteria used to judge whether a randomly generated population of galaxies is analogous to the observed HVG system. Only one of the first two criteria is used at a time. Note that the criteria are not all independent. For example, as the MW and M31 positions are fixed and my plane fitting procedure (Section 6.3.1) always considers them, the thinnest planes are likely to be obtained when these galaxies are close to the plane best fitting the HVGs. This makes it more likely that the 'barycentre offset' criterion is satisfied (compare corners of Table 10).

	Thickness	Direction	Barycentre offset
Thickness	4.6 ± 0.3		
Direction	2.3 ± 0.1	417.4 ± 0.5	
Barycentre offset	2.4 ± 0.2	81.1 ± 1.3	181.8 ± 2.5

Table 10. Monte Carlo trial-based probabilities in parts per thousand (‰) of the HVG system (all galaxies in Table 7) matching various combinations of the criteria defined in Table 9. These criteria are used to determine if a mock HVG system is analogous to the observed system, using the method outlined in Section 6.3. When the same criterion appears in both the row and column headings, the result is the probability of matching that criterion alone, regardless of the others. The probability of all three criteria being met simultaneously is 1.48 ± 0.10‰, which corresponds to the first row of Table 11.

on comparing different runs). In all cases, I determined the proportion of 'successful' MC trials to within a few percent.

The direction criterion was met in $\sim$417‰ of the trials and was the least problematic criterion. This is due to the rather wide range of orientations allowed for the plane best fitting the mock galaxies.

The plane of HVGs is offset from the MW-M31 barycentre by 67 kpc, which is rather small considering the extent of the HVG plane ($\sim$1 Mpc). Thus, the 'barycentre offset' criterion in Table 9 is only met around 182‰ of the time.

By far the most important criterion is the requirement that all but one mock HVG define a plane with rms thickness smaller than observed. This criterion is met in only 5.2 ± 0.2‰ of the MC trials. Consequently, it is very unlikely (probability 1.48 ± 0.10‰) that all three criteria are satisfied simultaneously. This remains a very unlikely situation despite various changes to my modelling assumptions and choice of sample, for example if I change the anisotropy statistic (Table 11). Neglecting distance uncertainties altogether has little impact on the results, suggesting that they should be robust to future improvements in distance measurements. The most significant change occurs if Antlia is removed from my HVG sample as it could be a satellite of NGC 3109 (van den Bergh 1999). However, even this case yields a very low probability of 5.4‰.

Investigation	Sample	Probability (‰)
Nominal (physical thickness)	All	1.48 ± 0.10
$\widehat{\boldsymbol{h}}$ rotated 5° south ($\theta = 75°$)	All	1.51 ± 0.10
Distances fixed	All	1.45 ± 0.01
Nominal	~~HIZSS 3~~	0.41 ± 0.02
Nominal	~~Antlia~~	5.17 ± 0.36
Aspect ratio	All	1.62 ± 0.01
Aspect ratio	~~Antlia~~	5.35 ± 0.02

Table 11. How my results depend on various model assumptions. The final column shows the probability of a MC trial satisfying the criteria given in Table 9 based on randomising the directions towards the HVGs in Table 7 but with a fixed M31 direction. Galaxies whose names have been crossed out are excluded from the sample in that particular investigation, with the nominal sample corresponding to the central column of Table 8. The exclusion of HIZSS 3 is achieved by altering Equation 70 to impose the requirement that any observable galaxy be $\geqslant 15°$ from the Galactic plane.

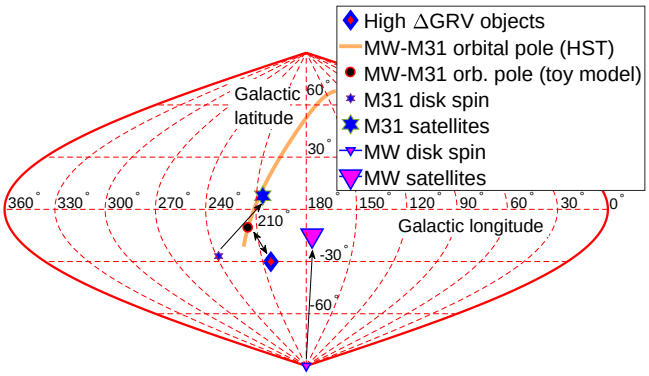


Figure 25. Normal directions to the important LG planes, shown in Galactic co-ordinates. Assuming a past close MW-M31 flyby, I expect tidal torque from M31 to explain the misalignment between the orientation of the MW disk (small triangle) and its plane of satellites (large triangle). The effect of such torques is illustrated with an upward arrow. Tidal torque from the MW explains a similar misalignment for M31 (hexagrams used instead of triangles). My MOND-based toy model is able to reproduce these orientations fairly well if the MW-M31 orbital pole lies in the direction of the black dot with red rim (Banik & Zhao 2018b, section 2.2). This is reasonably consistent with the normal to the plane defined by the HVGs (diamond), though its sense of rotation is unknown. The proper motion of M31 has recently been measured (van der Marel et al. 2012b), suggesting a particular MW-M31 orbital pole (1σ allowed region shown as orange line). This must be orthogonal to the present direction towards M31. Unfortunately, at 2σ, any direction consistent with this requirement is allowed.

Figure 25 shows how the HVG plane I found fits into other LG structures. If the HVG plane can be identified with the MW-M31 orbital plane, then this constrains models of a past MW-M31 flyby attempting to match the observed orientations of their satellite planes (e.g. Bílek et al. 2017).

To conclude this section, the LG has several galaxies with unusually high radial velocities in the context of ΛCDM. The spatial distribution of these HVGs is inconsistent with isotropy because they lie rather close to a well-defined plane. Thus, the spatial arrangement of the HVGs

is similar to what would be expected in the MOND scenario where their unusual kinematics arose due to past gravitational slingshot interactions with the MW/M31 around the time of the MW-M31 flyby. In this scenario, the anisotropy arose because the LG dwarfs flung out at the highest speeds were those flung out nearly parallel to the motion of the perturber. If this is correct, the MW-M31 line should lie within the HVG plane. In reality, it is only 16° off the plane, a possible consequence of tides raised by large scale structure and other effects not included in my simplified MOND model of the LG.

7 THE ESCAPE VELOCITY CURVE OF THE MILKY WAY IN MODIFIED NEWTONIAN DYNAMICS (Banik & Zhao 2018a)

So far, I have used galaxies merely as tracers for a timing argument analysis or as light sources which are gravitationally lensed. This is because there is already an extensive literature on testing MOND with forces internal to galaxies using their rotation curves (e.g. Swaters et al. 2009; Famaey & McGaugh 2012; Papastergis et al. 2016). It is difficult to analyse the forces within external galaxies much beyond measuring their radial acceleration profile in this way. However, additional measurements are possible for the MW because individual stars are resolved more easily, meaning their proper motions are often available (e.g. Zacharias et al. 2017) in addition to their radial velocities (e.g. Kunder et al. 2017).

In this section, I focus on the Galactic escape velocity curve. Although MOND is often assumed to imply a r^{-1} force law towards a point mass (Equation 2), this is only true if the mass is isolated. Even if more distant masses impose a constant external gravitational field on a system, the non-linearity of MOND implies that this affects the internal forces within the system. This external field effect (EFE) arises because the theory is acceleration-dependent (Milgrom 1986, section 2g). Ultimately, the EFE and the inherent non-linearity of MOND are required by data indicating that the force towards a galaxy of mass M scales more nearly as $\sqrt{M}$ rather than linearly with M (Figure 2).

To understand the EFE, consider a dwarf galaxy governed by MOND which has very low internal accelerations ($\ll a_0$) but is freely falling in the strong gravity ($\gg a_0$) of a distant massive galaxy. The overall acceleration at any point in the dwarf is rather high due to the dominant external field (EF) of the massive galaxy. Thus, the dwarf would obey Newtonian dynamics and forces in its vicinity would follow the usual inverse square law rather than Equation 2. This is true even if the EF is uniform across the dwarf i.e. there are no tidal effects. However, without the massive galaxy, the dwarf's internal dynamics would be very non-Newtonian.

Realising that MOND with the EFE predicts potential wells of finite depth, Famaey et al. (2007) used an analytic method to estimate the Galactic escape velocity v_{esc} from the Solar neighbourhood. Similar results were later obtained by Wu et al. (2008) using a numerical solution to AQUAL. Their estimated v_{esc} agrees reasonably well with later measurements based on high-velocity MW stars (Piffl et al. 2014). Recently, a similar technique was used to measure v_{esc} over a wide range of Galactocentric radii (8−50 kpc, Williams et al. 2017a). This work applied the method of Leonard & Tremaine (1990) to a variety of tracers detected in the ninth data release of the Sloan Digital Sky Survey (SDSS, Ahn et al. 2012). This section focuses on calculating the expected v_{esc} in MOND at these positions for a range of plausible Galactic mass models, bearing in mind constraints from the Galactic rotation curve (McGaugh 2016a).

7.1 Method

As in Section 6.2.1, the gravitational field of the MW is determined using Equation 59. The main complication is in calculating $\boldsymbol{g}_N$ accurately because the MW can no longer be treated as a point mass. The EF also has to be included. To avoid a total breakdown of symmetry in the problem, I assume the EF is aligned with the symmetry axis of the MW disk. The true EF direction may well be different, but this is expected to have only a very small effect on the results (Banik & Zhao 2018a, section 4.3).

The MW is assumed to consist of a hot gas corona surrounding two aligned and concentric infinitely thin exponential disks representing its gas and stellar components. Taking advantage of the fact that potentials superpose in Newtonian gravity, I simply add the potential of the corona to that of the other components. The corona is treated as a Plummer model (Plummer 1911) with mass M_{cor} and core radius r_{cor}, yielding a corona potential at a Galactocentric distance r of

$$\Phi_{cor} = -\frac{GM_{cor}}{\sqrt{r^2+{r_{cor}}^2}} \tag{73}$$

For the disk components, the superposition principle means that it is only necessary to solve for a single exponential disk. I take this to have unit scale length and GM, scaling it up to the required values later. To determine the Newtonian potential Φ_N of this mass distribution $\rho_b(\boldsymbol{r})$, I numerically solve the Poisson equation

$$\nabla^2\Phi_N = 4\pi G\rho_b \tag{74}$$

This is done using successive over-relaxation in spherical polar co-ordinates (polar angle θ). Further details are provided in appendix A of Banik & Zhao (2018a), which explains the discretisation scheme and convergence criteria.

To include the EF, I add the contribution from the Newtonian EF $\boldsymbol{g}_{N,ext}$. This is what the EF would have been in Newtonian gravity. I assume the spherically symmetric MOND relation between it and the actual EF $\boldsymbol{g}_{ext}$.

$$\Phi_N \rightarrow \Phi_N - \boldsymbol{r}\cdot\boldsymbol{g}_{N,ext} \qquad \text{where} \tag{75}$$

$$\overbrace{\nu\left(\frac{g_{N,ext}}{a_0}\right)}^{\nu_{ext}}\boldsymbol{g}_{N,ext} = \boldsymbol{g}_{ext} \tag{76}$$

Once all component potentials have been appropriately scaled according to the parameters in Table 12, it is easy to add them and thereby determine $\boldsymbol{g}_N = -\nabla\Phi_N$. As in Section 6.2.1, I then determine $\nabla\cdot\boldsymbol{g}$ using Equation 58 and apply direct summation (Equation 59) to obtain $\boldsymbol{g}$ at positions of interest. Some small corrections are then applied for edge effects (Banik & Zhao 2018b, section 2.2) using analytic approximations that are asymptotically correct (Banik & Zhao 2015b). I found the escape velocity v_{esc} by integrating the radial component of $\boldsymbol{g}$ along a radial transect

Variable	Meaning	Value
$R_\odot$	Galactocentric distance of Sun	8.2 kpc
r_*	Stellar disk scale length	2.15 kpc
$M_{*,0}$	Nominal stellar disk mass	$5.51 \times 10^{10} M_\odot$
r_g	Gas disk scale length	7 kpc
$M_{g,0}$	Nominal gas disk mass	$1.18 \times 10^{10} M_\odot$
$\frac{M_*}{M_{*,0}}$	Disk mass scaling factor	$0.8 - 1.4$
r_{cor}	Plummer radius of corona	$(20 - 60)$ kpc
M_{cor}	Corona mass	$(2 - 8) \times 10^{10} M_\odot$
g_{ext}	External field on MW	$(0.01 - 0.03)\, a_0$

Table 12. Parameters of the MW mass distribution, with $_0$ subscripts indicating nominal values while $_*$ and $_g$ subscripts refer to its stellar and gas components, respectively. I always use the same value of $\frac{M_g}{M_*}$ and the same disk scale lengths, but vary the other parameters using a grid search. The first part of the table contains the fixed parameters $R_\odot$ (McMillan 2017), $M_{*,0}$ (McGaugh 2016a), r_* (Bovy & Rix 2013), r_g (McMillan 2017) and $M_{g,0}$, which is based on applying the method described in McGaugh (2008, section 3.3) to the observations of Olling & Merrifield (2001, table D1). The Galactic hot gas corona is modelled using Equation 73 (Plummer 1911).

out to infinity. The nominal MW model used is designed to be consistent with its observed rotation curve in a MOND context (McGaugh 2016a, table 1 model Q4ZB), though I also consider several hundred other models.

7.2 Results

I begin by showing the circular and escape velocity curves of the MW with my nominal values for the MW stellar and gas disk masses (Figure 26). For comparison with observations, I fit a power-law model to v_{esc} over the radial range 10−50 kpc. This assumes that

$$v_{esc}\,(r) \;\propto\; r^{-\alpha} \tag{77}$$

Power-law fits become linear when considering the logarithms of both variables. Thus, if $y \equiv Ln\ v_{esc}$ and $x \equiv Ln\ r$ are lists of size N, then

$$\alpha \;=\; -\frac{\sum_{i=1}^{N} \tilde{x}_i \tilde{y}_i}{\sum_{i=1}^{N} \tilde{x}_i \tilde{x}_i} \quad \text{where} \tag{78}$$

$$\tilde{x} \;\equiv\; x - \frac{1}{N}\sum_{i=1}^{N} x_i \quad (\tilde{y} \text{ defined analogously}) \tag{79}$$

$v_{esc}\,(r)$ can be described rather well as a power law over the range $r = 10-50$ kpc (Figure 26). For a range of models, I determine the best-fit slope α and Solar circle normalisation for comparison with observations. It is unclear exactly which Galactic polar angles θ the observations of Williams et al. (2017a) correspond to, but most likely a range of angles is used in order to get enough of the relatively rare high-velocity stars that are necessary for an escape velocity determination. I show results within the disk plane ($\theta = \frac{\pi}{2}$) in Figure 27.

$v_{c,\odot}$ depends mainly on the disk surface density such that only the nominal value is able to correctly reproduce the observed LSR speed of $v_{c,\odot} \approx 235$ km/s (McMillan 2017). However, I also consider the effect of scaling the surface density by factors of 0.8−1.4. Raising this factor by 0.1 increases $v_{c,\odot}$ by ~ 8 km/s.

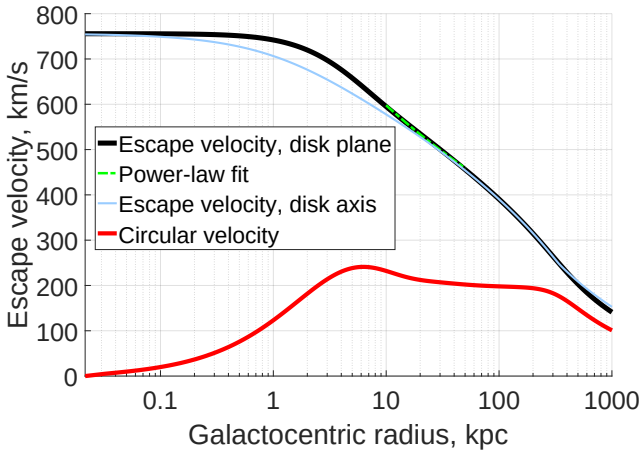


Figure 26. How the circular velocity of the MW (lower red curve) and its escape velocity (upper black curve) depend on position within its disk plane. The latter can be parametrised rather well as a power law (Equation 77) over the radial range $10 - 50$ kpc (dashed green curve). At the same distance from the MW, its escape velocity is lower along its disk axis (thin blue curve) for points close to the MW due to the effect of its disk. However, this pattern is reversed at long range because the EF on the MW is assumed to align with its disk axis, deepening the potential in this direction (Banik & Zhao 2015b, equation 37). The model shown here uses the nominal disk masses in Table 12 and $g_{ext} = 0.03 a_0$, with the corona being as small and low-mass as possible.

Within the range considered, adjusting M_{cor} affects $v_{c,\odot}$ by $\lesssim 5$ km/s while adjusting r_{cor} has a smaller effect of ~ 1 km/s. At the Solar circle, the MW is effectively isolated − adjusting g_{ext} only affects $v_{c,\odot}$ by ~ 4 m/s. These factors are more significant further from the MW, but the scarcity of tracers makes it difficult to directly measure $\boldsymbol{g}$ there. Fortunately, forces at large r affect the escape velocity $v_{esc} = \sqrt{-2\Phi}$ near the Sun. A local v_{esc} measurement could thus constrain the MW gravitational field at large distances, with the appropriate analysis. One of this section's objectives is to do just that, in a MOND context.

At the same r, escape velocities are slightly larger within the MW disk plane as the MW matter distribution is concentrated towards this plane. Within ~ 100 kpc of the MW, this near-field effect is more important than the non-sphericity of the MW potential in the far-field EF-dominated region, where the MW exerts very little gravity in any case. However, beyond ~ 100 kpc, the latter effect dominates because the MW can be considered as a point mass (compare black and blue curves in Figure 26). As shown by Banik & Zhao (2015b), this leads to a deeper potential in the direction of $\boldsymbol{g}_{ext}$ i.e. along the disk axis in my axisymmetric models.

An important constraint on the true Galactic $v_{esc}\,(r)$ curve arises because it is determined by the same potential Φ that governs the rotation curve $v_c\,(r)$.

$$\frac{\partial\left(\frac{1}{2}{v_{esc}}^2\right)}{\partial r} \;=\; -\frac{\partial \Phi}{\partial r} \;=\; -\frac{{v_c}^2}{r} \tag{80}$$

$$\alpha \;\equiv\; -\frac{\partial Ln\ v_{esc}}{\partial Ln\ r} \;=\; \left(\frac{v_c}{v_{esc}}\right)^2 \tag{81}$$

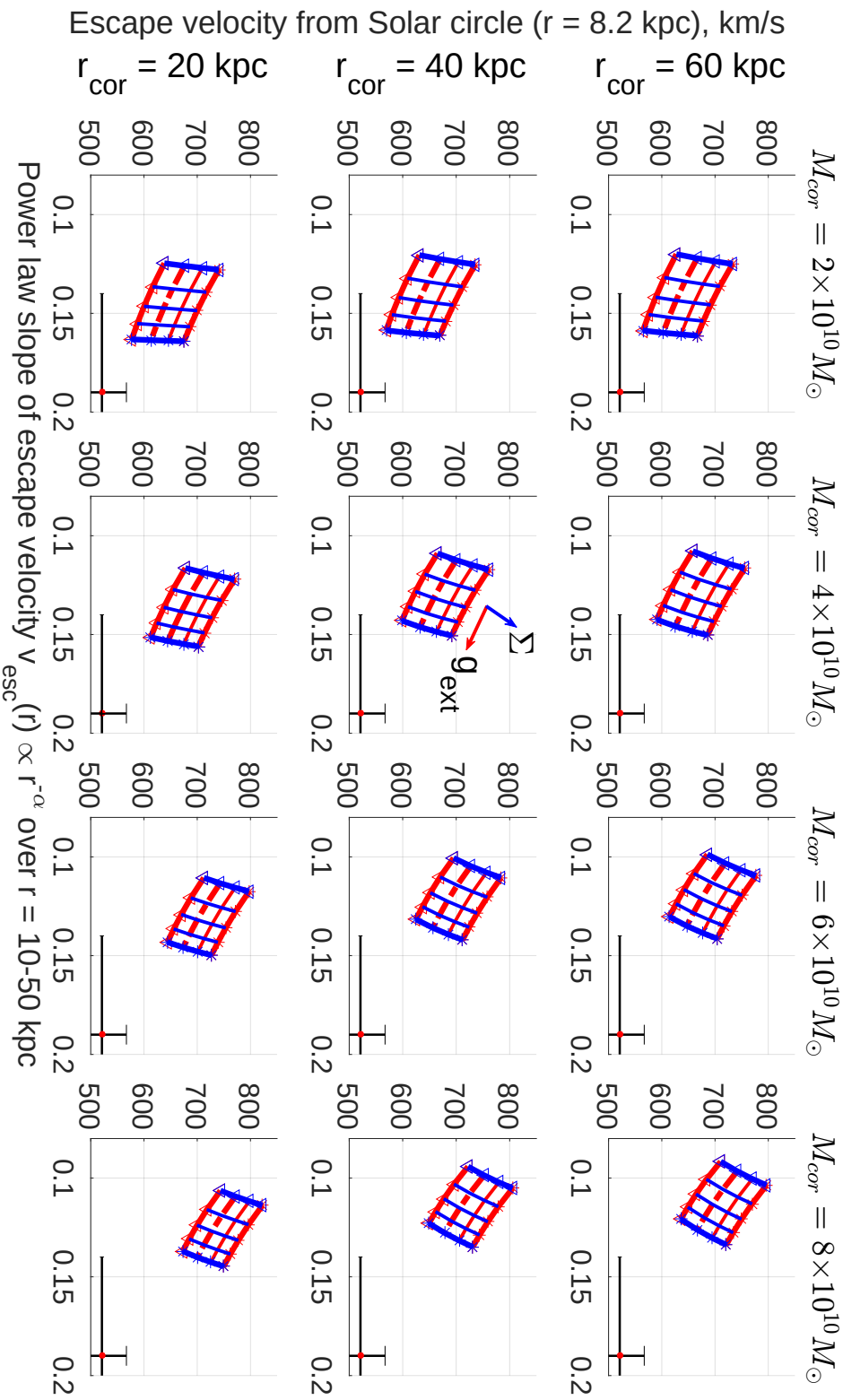


Figure 27. Escape velocity v_{esc} from within the MW disk plane as a function of model parameters. The x-axis shows the value of α such that $v_{esc}(r) \propto r^{-\alpha}$ while the y-axis shows v_{esc} near the Sun. The measured values of these quantities are shown as a red dot with black error bars towards the bottom right (Williams et al. 2017a). Each subplot has a fixed corona mass and scale length, with red tracks showing the effect of varying g_{ext} with constant disk mass (vice versa for blue tracks). In each case, an inverted triangle is used to show the result when the parameter being varied has the lowest value considered while a star is used for the largest value. This is also shown by the arrows in the central subplot, which point towards higher values of the indicated parameter. I consider disk masses scaled from the nominal value by factors given in Table 12, where I also show the range in g_{ext} that I try (values of all parameters are spaced linearly). The dashed red lines show the results for the nominal stellar and gas disk masses, which is required to obtain the correct $v_{c,\odot}$ (McGaugh 2016a, table 1 model Q4ZB). I assume $\boldsymbol{g}_{ext}$ is aligned with the disk symmetry axis. Rotate 90° anti-clockwise for viewing.

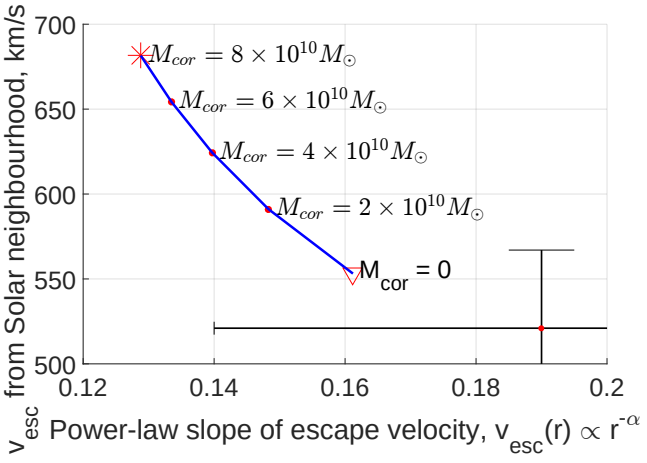


Figure 28. Effect of the MW corona mass M_{cor} on its escape velocity curve for points along its disk axis in the direction of the external field. Other model parameters are the same as in Figure 26. The x-axis shows the value of α such that $v_{esc}(r) \propto r^{-\alpha}$ while the y-axis shows v_{esc} near the Sun. The measured values of these quantities are shown as a red dot with black error bars towards the bottom right (Williams et al. 2017a).

If $v_{c,\odot} = 232.8$ km/s (McMillan 2017), then $\alpha = 0.200$ for a local escape velocity of 521 km/s. This is entirely consistent with the observed value of 0.19 ± 0.05. It is clear that my calculated escape velocities are towards the upper end of the range allowed by observations. Thus, my analysis disfavours a hot gas corona. I have included one because XMM-Newton (Jansen et al. 2001) observations at a range of Galactic latitudes indicate that one is present (Nicastro et al. 2016). Their best-fitting model suggested that its mass is $2 \times 10^{10} M_\odot$ (see their table 2 model A) which is therefore the lowest value for M_{cor} that I consider. Similarly to my analysis, the best fit to their observations was obtained for the lowest mass corona model they tried out, though substantially more massive halos are far from ruled out.

A hot gas corona would also cause ram pressure stripping effects on MW satellites containing gas. This is thought to explain the asymmetry of the Magellanic Stream (Hammer et al. 2015) and perhaps also the truncation of the LMC gas disk at a much shorter distance than the extent of its stellar disk (Salem et al. 2015). Those authors used this argument to estimate that $M_{cor} = 2.7 \pm 1.4 \times 10^{10} M_\odot$, consistent with other estimates.

Although my analysis is consistent with this, it prefers an even lower M_{cor}. I therefore considered lowering M_{cor} all the way down to 0. As expected, this makes the MW v_{esc} curve slightly more consistent with observations in terms of both its amplitude and its radial gradient (Figure 28).

The Galactic escape velocity curve (Williams et al. 2017a) is consistent with expectations in MOND based on a MW mass model that also explains its rotation curve (McGaugh 2016a, table 1 model Q4ZB). A fairly low mass corona is preferred, consistent with independent measurements (Nicastro et al. 2016, table 2 model A). It is presently difficult to use the Galactic v_{esc} curve to meaningfully constrain how extended its corona is. Within the range considered, my analysis prefers a strong EF with $g_{ext} = 0.03 a_0$, a value also assumed by Famaey et al. (2007).

8 FUTURE PROSPECTS

MOND represents a significant departure from Newtonian dynamics. This should allow definitive tests in the near future. In this section, I discuss some possible ways of distinguishing the theories at a variety of astrophysical scales.

8.1 Beyond the Local Group

In Section 3, I discussed how measurements of the Moving Cluster Effect could help determine the collision velocity of interacting galaxy clusters like the Bullet Cluster and El Gordo (Menanteau et al. 2012). Much faster collision velocities are expected in MOND (Candlish 2016). Confirmation of the high estimated collision velocity in the Bullet Cluster and discovery of even a small number of other similar systems could severely challenge ΛCDM (Kraljic & Sarkar 2015).

In this paradigm, close interactions between galaxies rapidly end in a merger (Privon et al. 2013) due to dynamical friction (Chandrasekhar 1943) between overlapping DM halos. Without these halos, merger rates are expected to be much lower. Unfortunately, it is difficult to test this directly because there is a degeneracy between the frequency and visible duration of galactic interactions. Merger rates could be constrained through gravitational waves (GWs) emitted from merging supermassive black holes (SMBHs). Null detections using pulsar timing arrays are in some tension with the expected frequency of such events (Shannon et al. 2015).

The expected GW background is not expected to differ much depending on precisely how the progenitor black holes eventually merge, a question often called the 'final parsec problem'. Either this occurs through dynamical friction against stars and gas, or this process becomes inefficient near the final parsec. In this case, SMBHs from several progenitor galaxies would orbit within a rather small region such that dynamical friction would effectively arise against this population of objects − SMBHs would interact with each other (Ryu et al. 2018). Moreover, those authors showed that a longer merging timescale would mean mergers typically occurred later and thus closer to the Earth, making for a stronger GW here. This nearly cancels the effect of fewer mergers occurring altogether (see their section 4.1.1). In any case, the GW background due to merging SMBHs ought to become detectable in the near future (Wang & Mohanty 2017), thereby constraining cosmological models. Indeed, existing measurements are already placing interesting constraints (NANOGrav Collaboration 2018).

As well as detecting these GWs directly, the momentum they carry could be detected indirectly because it ought to cause the remnant SMBH to recoil. Given that the first GWs to be detected carried off $\sim 5\%$ of the progenitor's rest mass as energy (LIGO Collaboration 2016), the recoil could be significant if the GW emission is even slightly asymmetric. This would cause the SMBH to oscillate in the potential of its host galaxy, leading to an offset between the photometric centres of nearby elliptical galaxies and the positions of their central SMBH (often identifiable as an active galactic nucleus). Lena et al. (2014) searched for these offsets but found only small ($\lesssim 10$ pc) offsets. As frequent mergers are expected in ΛCDM, larger offsets should have been detected in some of the 14 cases considered. Moreover, even the small detected

offsets were often aligned with the jet created by accretion onto the SMBH, suggesting that hemispherical asymmetries in its power are responsible for the observed offset (see their section 5.3).

Extending the analysis to more galaxies should give a better statistical understanding of whether significant SMBH-host galaxy offsets are common. More detailed modelling is required to understand what this implies about the merger rate of galaxies. If it is very difficult for the progenitor SMBHs to inspiral sufficiently for GW emission to become significant, then there should be galaxies with multiple 'stalled' SMBHs near their centre that could ultimately be revealed through detailed multi-epoch kinematic measurements (e.g. Wang et al. 2017). If instead the binary SMBH orbit ought to rapidly decay, then it will be important to understand whether the eventual merger is likely to cause asymmetric emission of GWs and thus a detectable recoil. Without a convincing explanation for why the GWs ought to be symmetric, continued null detection of large SMBH-galaxy offsets would strongly suggest a low major merger rate (depending on how efficiently the oscillations are expected to be damped).

Reduced dynamical friction between galaxies would allow close interactions at a much higher relative velocity. This could be tested by searching for galaxies with a high pairwise relative velocity but a small separation, such as might be the case for NGC 1400 and NGC 1407 (Tully et al. 2013; Tully 2015) and perhaps also for NGC 6050 and IC 1179. A small physical separation might be discernible from tidal features connecting the galaxies. This could also help prove that they were observed past pericentre, when dynamical friction between their DM halos should have slowed them down.

As well as the dynamics of the interacting galaxies themselves, an important issue is the properties of any tidal dwarf galaxies (TDGs) that form out of the encounter. As discussed in Section 1.6, it is critical to understand whether the acceleration discrepancy persists in such systems as it should not in ΛCDM (Wetzstein et al. 2007). This is precisely what was investigated by Gentile et al. (2007) based on observations of the NGC 5291 system (Bournaud et al. 2007). However, it was later realised that these TDGs formed rather recently, leaving them insufficient time to settle into dynamical equilibrium (Flores et al. 2016). Even so, detailed observations of much older TDGs remain a promising way to understand how the acceleration discrepancies arise. One possible target is the ~ 4 Gyr old TDG identified by Duc et al. (2014), which may well have settled into dynamical equilibrium by now. Another very promising set of targets are the members of the recently discovered satellite plane around Cen A (Müller et al. 2018).

8.2 Within the Local Group

Although some systems outside the LG would behave very differently in MOND and ΛCDM, the large distance to these systems makes it difficult to tell whether this is actually the case. This is why most of my thesis has focused on the LG, even though any signatures of MOND are likely to be more subtle.

One obvious way to test MOND within the LG is to simulate the MW-M31 flyby in more detail using N-body and (eventually) hydrodynamic models that incorporate MOND gravity. This is feasible using the Phantom of RAMSES algorithm (Lüghausen et al. 2015), an adaptation of the RAMSES algorithm widely used by astronomers (Teyssier 2002). Precisely this sort of simulation has recently been done (Bílek et al. 2017). It is important to search the parameter space more thoroughly to see if some model can explain the observed orientations of the MW and M31 satellite planes as arising from tidal debris expelled during their interaction.

I recently conducted an investigation along these lines with the help of a summer student I hired (Banik et al. 2018). Treating the MW and M31 as point masses, I considered a disk of test particles around each one and advanced their trajectories using a very similar method to that in Section 6.2. I tried out a range of MW-M31 orbital poles consistent with the small observed proper motion of M31 (van der Marel et al. 2012b). Unfortunately, this does not reliably constrain the orbital pole much beyond the fact that it must be orthogonal to the present direction in which we observe M31. Thus, I tried the full range of allowed directions and a range of MW-M31 tangential speeds, leading to a range of closest approach distances.

In each model, I looked at the distribution of tidal debris outside the disk plane of each galaxy but within 250 kpc. The orbital poles of particles in this 'satellite region' are shown in Figure 29, with each particle statistically weighted according to the mass it represents within the disk it originated in. Around each galaxy, the orbital poles show a clear clustering, as happens in most models. In the particular one shown, the preferred orbital poles align fairly well with the observed orientations of the MW and M31 satellite planes. Around the MW, the model also gets some material on orbits that are roughly counter-rotating with respect to the preferred rotation direction. This is interesting as the MW satellite Sculptor is counter-rotating within the MW satellite plane (Sohn et al. 2017).

My model does not get counter-rotators around M31, a prediction that could be tested with proper motions of its satellites within its satellite plane. In particular, the radial velocities of And XIII and And XXVII suggest that they may be counter-rotators. However, they could just be interlopers whose orbits take them far from the satellite plane. This seems quite feasible given that only about half of the M31 satellites lie within its satellite plane structure (Ibata et al. 2013).

As discussed further in Banik et al. (2018), the model also gets a reasonable radial distribution of tidal debris and thickens the MW disk by a similar amount to its observed thickness. The LG mass needed to satisfy the timing argument (Equation 60) is similar to what the MW and M31 rotation curves imply, without assuming significant mass loss during their flyby. This occurs roughly when observations indicate the MW thick disk formed (Quillen & Garnett 2001). However, the MW-M31 orbital pole in the model is nearly 60° from the HVG plane identified in Section 6. This may be because the HVGs identified there are mostly from a single bound association that has now disrupted after a close passage with the MW or M31. Such a common origin would naturally explain the filamentary nature of the NGC 3109 association (Bellazzini et al. 2013). Thus, the orientation of the HVG plane might not be securely determined yet.

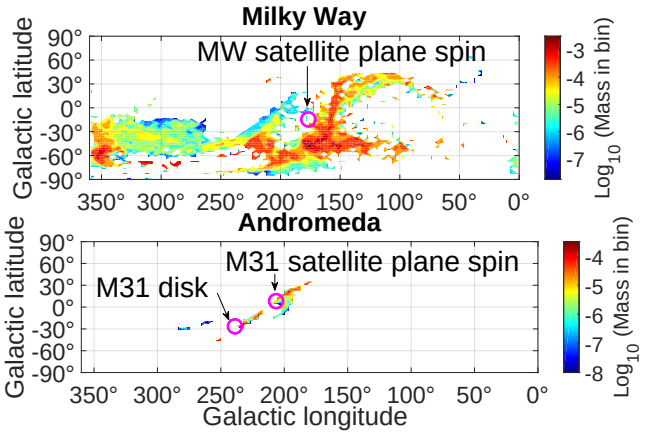


Figure 29. The distribution of orbital angular momentum directions (spin vectors) for tidal debris around the MW and M31 disks at the end of my best-fitting restricted N-body simulation of a past MW-M31 flyby (Banik et al. 2018). The mass units are arbitrary. *Top*: Results for the MW. Its disk spin vector points at the South Galactic Pole while the open pink circle shows the spin vector of its satellite plane. *Bottom*: Results for M31. I use open pink circles to show the observed spin vectors of its disk (lower left) and satellite plane (upper right).

In addition to a past MW-M31 flyby, MOND also has more subtle consequences in the LG. One very interesting example is the EFE – the internal dynamics of a system should be affected by the constant EF in which it is embedded, even in the absence of tidal effects (e.g. Banik & Zhao 2015b). This violates the strong equivalence principle. Perhaps the most accurate current test of the EFE is the internal velocity dispersion σ of the recently discovered MW satellite Crater 2 (Torrealba et al. 2016). Without the EFE, σ should have been $\sim$4 km/s in MOND but including the EFE (which is natural to MOND) reduces this to 2.1 km/s (McGaugh 2016b), mainly because its internal accelerations are reduced by its rather large half-light radius of 1066 ± 84 pc (Torrealba et al. 2016). σ was later observed to be 2.7 ± 0.3 km/s (Caldwell et al. 2017), a major topic of discussion at the Cleveland debate between MOND and ΛCDM.

In ΛCDM, large satellites like this ought to probe a significant part of their DM halo. This makes it difficult to argue that the visible extent of Crater 2 only probes the rising part of its rotation curve. As the baryon fraction needs to be very low in low mass DM halos to explain their internal dynamics, they must retain only a very small fraction of their baryons. However, if an object with such a low σ can retain visible baryons at all, then there ought to be many more satellites with slightly higher σ. This would worsen the missing satellites problem whereby the MW satellite mass function does not match the distribution of DM subhalo masses expected in ΛCDM (Klypin et al. 1999).

This issue could be resolved if the DM halo of Crater 2 was tidally stripped during close passage(s) with the MW (Fattahi et al. 2017). This is possible if Crater 2 is on a very eccentric Galactic orbit. Otherwise, it might be difficult for ΛCDM to explain its very low internal velocity dispersion for its size while also remaining consistent with statistical properties of MW satellites. The orbital history of Crater 2 should become much clearer once its proper motion is known, making this an important test of the ΛCDM paradigm. Thus, it is fortunate that Sohn (2016) proposed taking this measurement.

8.2.1 Within the Milky Way

Several detailed tests of MOND should become possible with MW data collected by the GAIA mission (Perryman et al. 2001). One of these is based on the vertical force towards the MW disk at a range of Galactocentric radii (Bienaymé et al. 2009). MOND predicts that the vertical force is boosted by the local factor of ν, which is radius-dependent. Thus, the Newtonian dynamical disk surface density should decline outwards in a different way to that of the visible MW baryons. Another test is based on accurately measuring the shape of the stellar velocity dispersion tensor several kpc from the Galactic disk plane, possibly near the Solar Circle.

Due to an effect similar to the MCE (Section 3), some constraints can be placed based on the plethora of DM substructure that ought to arise in ΛCDM. The motion of a DM mini-halo between us and a pulsar would occasionally cause the observed period of the pulsar to decrease if such halos were sufficiently common. As pulsar periods generally increase, this has allowed interesting constraints to be placed on low-mass DM halos (Clark et al. 2016). Future pulsar timing observations could improve these constraints further, regardless of assumptions concerning whether DM undergoes self-annihilation. The precise particle nature of the DM is still somewhat relevant because if it is of a sufficiently low mass, then it would not form very low mass halos (Viel et al. 2013). However, such warm DM scenarios are strongly constrained by Lyman-α forest data which reveals plenty of very low mass gas clouds that likely delineate low mass DM halos (Iršič et al. 2017).

MOND is an acceleration-dependent modification to gravity, so the transition from Newtonian to modified dynamics can arise at a much smaller length scale than the sizes of galaxies whose rotation curves originally motivated the theory. This is possible if one focuses on much lower mass systems. One such situation that has recently attracted some attention is wide binary stars (Hernandez et al. 2012). This is based on the MOND radius of the Sun being only 7000 AU (Equation 2), so two Sun-like stars separated by this distance or more should rotate around each other faster than expected in Newtonian gravity. In general, the actual rotation speed could exceed the Newtonian expectation by an arbitrarily large factor. However, in the Solar neighbourhood, the 'external' gravitational field from the rest of the Galaxy limits the boost to gravity that could be provided by MOND. With the simple interpolating function (Famaey & Binney 2005), this boost is up to a factor of 1.56.

Such an enhancement to the self-gravity of wide binaries should become detectable in the GAIA era (Scarpa et al. 2017). This issue was recently investigated in some detail by Pittordis & Sutherland (2017), who showed that the prior distribution of orbital eccentricity and semi-major axis should not much affect the conclusions. Unbound wide binaries would disperse rather quickly – at typical orbital velocities of $\sim \sqrt{\frac{GM_\odot}{7000\,AU}} = 0.3$ km/s, the separation should rise to 1 pc in only a few Myr. But wide binaries in the

MOND sense are still much closer together than field stars in the MW, making for only a very small chance that two unbound stars would randomly be so close together in 3D. Also requiring a similar 3D velocity would reduce the contamination further, making the wide binary test a compelling way of constraining what law of gravity governs gravitational systems with accelerations typical of galactic outskirts.

The wide binary systems necessary for this test are expected to be quite common (Andrews et al. 2017). In fact, our nearest external star system consists of Proxima Centauri (Proxima Cen) orbiting the close binary α Cen A and B at a distance of 13000 AU (Kervella et al. 2017). This puts the Proxima Cen orbit well within the regime where MOND would have a significant effect (Beech 2009, 2011). As well as suggesting that wide binaries ought to be common, even this single system could allow a direct test of MOND with the proposed Theia mission (Theia Collaboration 2017).

To see how this might work, I used an algorithm similar to that in Section 6.2.1 to model the orbit of Proxima Cen. I treated it as a test particle orbiting the much more massive α Cen A and B, which I considered as a single point mass of $2.043M_\odot$ given that they are in a tight orbit separated by only ~ 18 AU (Kervella et al. 2016). The EF was taken to be towards the Galactic centre and of a magnitude sufficient to maintain the observed $v_{c,\odot}$ of 232.8 km/s, assuming the Sun is 8.2 kpc from the Galactic centre (McMillan 2017). As this is much larger than the distance to α Cen, it feels nearly the same $\boldsymbol{g}_{ext}$ as the Sun.

I used the gravitational field found in this way to integrate the orbit of Proxima Cen forwards, starting with the radial velocity and proper motion measurements in table 2 of Kervella et al. (2017). I also found the Proxima Cen trajectory in Newtonian gravity. In both cases, the observations are assumed to span a negligibly short fraction of the ~ 500 kyr orbital period, allowing the force to be approximated as constant and the trajectory as parabolic.

The angular difference between the trajectories on our sky is shown in Figure 30. Unless the initial conditions are known exactly, the difference would actually be $\frac{1}{4}$ of that shown because astronomers would try to fit the data using different initial conditions.[1] Even so, a parabola can only be fit with a straight line for so long. Thus, if Theia is flown and achieves μas astrometric precision over a few years, it should be able to directly determine how much Proxima Cen accelerates towards α Cen A and B. This is because the acceleration is expected to be $\sim 40\%$ higher in MOND compared to Newtonian dynamics, though it is only $\sim a_0$ in either case and thus very small.

In principle, the radial velocity v_r of Proxima Cen could also be used to distinguish these theories. However, a constant acceleration causes v_r to change linearly with time, whereas the position would respond quadratically. Thus, v_r would only differ by 0.5 cm/s between the models after a decade of observations. This would be very challenging to detect, making it a much less plausible test of MOND than using precise astrometry of Proxima Cen.

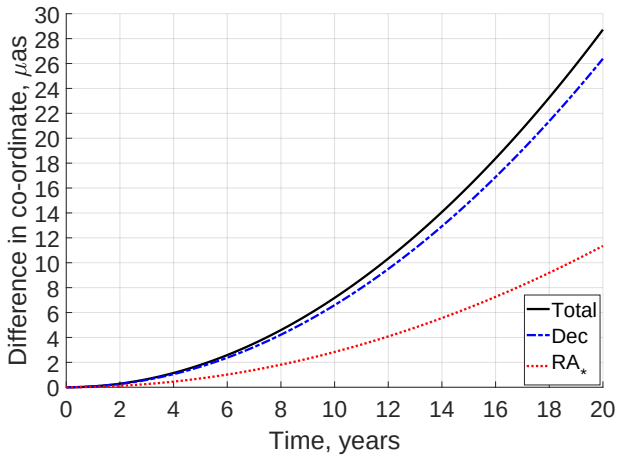


Figure 30. Difference in sky position of Proxima Cen depending on whether Newtonian gravity or MOND governs its orbit about α Cen A and B. The same initial conditions are used for both trajectories (Kervella et al. 2017, table 2). The difference in right ascension has been scaled by the cosine of the declination so it corresponds to an actual angular difference. The total angular difference grows quadratically with time and is 7.18 μas after 10 years. Astronomers might try to fit the data by varying the initial conditions, in which case the angular differences would be $\sim \frac{1}{4}$ that shown here.[1]

One possible complication with such tests is that an undetected exoplanet could also cause an extra acceleration. However, as perceived at Proxima Cen, the exoplanet is quite likely to be in a different direction than α Cen. Moreover, a short period exoplanet would show up in multi-epoch observations. This would not work with a sufficiently long period, but in this case the greater distance implies the exoplanet must be more massive and so more likely to be detected. This is especially true given our proximity to the system enlarging the angles involved, thus making it easier to achieve sufficient starlight suppression in the region of interest. If an anomalous acceleration was detected, then intensive observations could be taken in its direction from Proxima Cen.

In the long run, interstellar precursor missions should be able to test MOND directly as $\nu_{ext} \approx 1.5$ in the Solar neighbourhood due to the rest of the MW. Thus, a Cavendish-type experiment performed $\gtrsim 10$ kAU from the Sun should yield rather different forces in MOND than for a similar experiment on the Earth. Alternatively, laser ranging measurements to a spacecraft at such a distance could be used to directly constrain the gravity exerted on it by the Sun (Christian & Loeb 2017). The Breakthrough Starshot initiative plans to send spacecraft to much larger distances and thus explore the nearest stars (Merali 2016). Testing MOND may provide a valuable incentive for interstellar precursor missions that only reach a few percent of this distance.

For such tests to be accurate, the non-gravitational accelerations would need to be constrained as these were important in the case of the Pioneer anomaly – this probably arose due to anisotropy of radiation emitted by the spacecraft (Turyshev et al. 2012). As any on-board computations

[1] The exact ratio will depend on spacecraft performance and other factors. I assumed the fit to data would be designed to minimise its χ^2 with respect to observations taken at regular intervals with equal accuracy. In this case, the best linear fit to the parabola $y = x^2$ over the range $0 - 1$ is given by $y = \frac{3}{4}x$.

require energy to be radiated, such effects can't be completely avoided. Although they could perhaps be reduced, it is almost certainly necessary and quite feasible to have an on-board accelerometer that accurately measures such non-gravitational accelerations (Lenoir et al. 2011). Combined with tracking data, this could allow rather sensitive tests of gravity.

8.3 Extensions to the basic MOND paradigm

The MOND paradigm may need further modification to satisfy all observational constraints. It has been suggested that an additional sterile neutrino species with an 11 eV mass could resolve outstanding issues of MOND at the galaxy cluster scale (Angus et al. 2010) and with the cosmic microwave background radiation (Angus 2009). This hot DM model gets the large scale structure of the Universe approximately correct, but has difficulty matching the detailed shape of the matter power spectrum at galaxy cluster scales (Angus & Diaferio 2011). Interestingly, recent structure growth measurements reveal some tension with ΛCDM that may be reduced if at least some of its DM was dynamically hot (Nesseris et al. 2017).

Other hybrid MOND-DM approaches have also been considered, for instance the superfluid DM model (Khoury 2016). It suggests that galaxies are surrounded by DM halos which cause additional non-gravitational forces between the baryons mediated by phonons in the superfluid DM. This can naturally reproduce the RAR. Even purely baryonic satellites of the MW within its DM halo would behave as if they were governed by MOND. However, the large extent of the MW and M31 satellite planes means that the model faces similar issues to ΛCDM − the more distant satellite plane members would still be Newtonian. Moreover, superfluid DM only creates an extra non-gravitational force on baryons, making it unclear how the theory can explain galaxy-galaxy weak lensing measurements that are consistent with light deflection governed by the RAR but not by standard gravity of the baryons alone (Milgrom 2013). Although the normal gravitational effect of the superfluid DM would make it act similarly to ΛCDM on large scales, its distribution would need to be rather finely tuned to satisfy Equation 11 and avoid radii where the deflection angle undergoes a Keplerian decline before rising again due to the DM halo. In fact, this may be a problem for ΛCDM (e.g. Faber & Gallagher 1979), motivating some workers to consider adjusting the basic paradigm (e.g. Kamada et al. 2017).

Given these difficulties, it is entirely possible that the empirical MOND equations need further modification. As the cluster-scale issues faced by MOND (Sanders 2003) arise in regions with a deep potential well, it has been suggested that the gravitational potential affects the acceleration parameter a_0 (Zhao & Famaey 2012). This theory of extended MOND (EMOND) can be tuned to match galaxy and cluster-scale observations fairly well (Hodson & Zhao 2017). Further work could help determine if EMOND is consistent with the internal dynamics of galaxies in cluster environments, where the local value of a_0 should be higher than in more isolated field galaxies.

MOND is an empirical theory whose more fundamental basis (if any) could help reconcile issues faced by its current formulation. One approach to understanding why MOND arises at all is the emergent gravity theory (Verlinde 2017). Although it is underpinned by some interesting theoretical ideas, the model faces difficulty explaining the observed RAR (Lelli et al. 2017a). A decisive test of the theory should be possible in the near future because it predicts extra perihelion precession of Solar System planets, especially Mars (Iorio 2017).

Originally, MOND was formulated as a non-relativistic theory (Bekenstein & Milgrom 1984). It can be embedded within a relativistic framework (Bekenstein 2004), a very relevant exercise in light of recent GW detections (LIGO Collaboration 2016). Expecting these results, GW propagation in MOND was briefly explored by Milgrom (2014), though of course further work would be highly relevant. It is particularly important to check whether relativistic MOND theories are consistent with GWs propagating at speeds very close to the speed of light, as required to explain recent observations of GW170817 (Virgo & LIGO Collaborations 2017). Perhaps MOND can't be reconciled with these observations (Chesler & Loeb 2017). It is unclear whether their arguments rule out particular relativistic generalisations of MOND or the paradigm in its entirety. As MOND was originally formulated to deal with non-relativistic systems (galaxies), the former seems more likely. However, only time will tell if it is possible to formulate MOND in a way consistent with special relativity and the near-simultaneous arrival of GWs and their electromagnetic counterpart. Even if this is possible, the galaxy cluster and larger scale issues faced by MOND may yet prove its undoing.

9 CONCLUSIONS

In this thesis, I explored several situations within and beyond the Local Group (LG) that are likely to reveal behaviour characteristic of MOND if it is correct. I began by considering how galaxy cluster collision velocities could be measured more accurately to search for a high-velocity tail not expected in ΛCDM but expected if gravity is enhanced at long range (Section 3). Despite possible differences of several hundred km/s, the large distance to such systems makes it difficult to determine their kinematics. Thus, the remaining contributions in this portfolio focus on the LG.

In Section 4, I described a ΛCDM dynamical analysis of the LG that treated it as axisymmetric about the MW-M31 line. This revealed several galaxies with much higher radial velocities than expected. I suggested that these HVGs were flung out by the MW/M31 around the time of their past close flyby, an event which would have happened in MOND (Zhao et al. 2013) but not ΛCDM due to dynamical friction between their DM halos (Privon et al. 2013).

In Section 5, I used an algorithm provided by one of the founders of the ΛCDM paradigm (Peebles) to address the issue of whether it really faces difficulty explaining the kinematics of LG dwarfs. The main difference was that the algorithm used a full 3D model, allowing rigorous consideration of tides raised by external galaxies and galaxy groups within 10 Mpc. Despite using a different algorithm written by different people in different programming languages, my conclusions remained broadly similar and were confirmed in a similar analysis by Peebles (2017). Such HVGs do not easily arise in cosmological simulations due to interactions

of LG dwarfs with analogues of MW or M31 satellites (Sales et al. 2007, figures 3 and 6). Thus, although some real LG satellites may be missing from my model due to a lack of detectable baryons, it is unclear how this would explain my results as such satellites should still be present in ΛCDM simulations.

To test my proposed scenario regarding the HVGs, I focused on comparing their spatial distribution and other properties to what these ought to be if the HVGs really were flung out by three-body interactions with the MW and M31 (Section 6). I found that the HVGs should preferentially lie within the plane of the MW-M31 orbit (Figure 23). The HVGs do indeed define a rather thin plane (Figure 24) oriented so the MW-M31 line is only 16° out of this plane. Thus, the properties of the HVGs are broadly what they should be if their anomalous kinematics arose due to passing near the spacetime location of a past MW-M31 flyby.

To gain experience with detailed MW models of the sort likely to become testable in the GAIA era, I calculated the escape velocity curve of the MW in MOND (Section 7) and compared it with recent measurements covering Galactocentric radii of 8−50 kpc (Williams et al. 2017a). I was able to account for both the amplitude and radial gradient of the Galactic escape velocity curve using a baryonic mass model consistent with the much more accurately known MW rotation curve (McGaugh 2016a, table 1 model Q4ZB). My results suggest that the MW has only a fairly low mass hot gas corona around it and is embedded in an EF of $\sim 0.03a_0$.

In Section 8, I briefly reviewed some possible avenues for future investigations to help determine if there really are substantial deviations from standard gravity at low acceleration. Although many of these ideas could be useful, the wide binary test of gravity seems the most promising near-term line of attack due to the impending release of data from the GAIA mission (Perryman et al. 2001). This would be an almost direct test involving a very different type of low-acceleration system to the rotating disk galaxies which originally inspired MOND. Nonetheless, both systems are alike in that their overall density greatly exceeds the cosmic mean value. Therefore, the application of MOND to these non-relativistic systems is quite clear.

10 ACKNOWLEDGEMENTS

10.1 General acknowledgements

The work on the high-velocity galaxy plane (Section 6) was suggested by Marcel Pawlowski and benefited from a visit to Princeton hosted by Nima Arkani-Hamed at the Institute for Advanced Studies. That work and the 3D ΛCDM model of the Local Group (Section 5) were both based on a FORTRAN algorithm kindly lent by P. J. E. Peebles, who also provided much advice regarding its operation. All the other algorithms were set up using MATLAB®.

10.2 Funding

IB is supported by Science and Technology Facilities Council studentship 1506672. The visit to Princeton was funded by a Scottish Universities' Physics Alliance travel grant.

REFERENCES

ATLAS Collaboration 2015, Journal of High Energy Physics, 10, 134
Ahmed S. H., Brooks A. M., Christensen C. R., 2017, MNRAS, 466, 3119
Ahn C. P., et al., 2012, ApJS, 203, 21
Alcock C., et al., 2000, ApJ, 542, 281
Andrews J. J., Chanamé J., Agüeros M. A., 2017, MNRAS, 472, 675
Angus G. W., 2009, MNRAS, 394, 527
Angus G. W., Diaferio A., 2011, MNRAS, 417, 941
Angus G. W., Famaey B., Diaferio A., 2010, MNRAS, 402, 395
Angus G. W., Diaferio A., Kroupa P., 2011, MNRAS, 416, 1401
Aragon-Calvo M. A., Silk J., Szalay A. S., 2011, MNRAS, 415, L16
Babcock H. W., 1939, Lick Observatory Bulletin, 19, 41
Baer H., Choi K.-Y., Kim J. E., Roszkowski L., 2015, Physics Reports, 555, 1
Banik I., 2014, preprint, Arxiv (arXiv:1406.4538v2)
Banik I., Zhao H., 2015a, MNRAS, 450, 3155
Banik I., Zhao H., 2015b, preprint, Arxiv (arXiv:1509.08457)
Banik I., Zhao H., 2016, MNRAS, 459, 2237
Banik I., Zhao H., 2017, MNRAS, 467, 2180
Banik I., Zhao H., 2018a, MNRAS, 473, 419
Banik I., Zhao H., 2018b, MNRAS, 473, 4033
Banik I., O'Ryan D., Zhao H., 2018, MNRAS, accepted
Barnes J. E., Hernquist L., 1992, Nature, 360, 715
Beech M., 2009, MNRAS, 399, L21
Beech M., 2011, ApSS, 333, 419
Bekenstein J. D., 2004, Physical Review D, 70, 083509
Bekenstein J., Milgrom M., 1984, ApJ, 286, 7
Bell E. F., de Jong R. S., 2001, ApJ, 550, 212
Bellazzini M., Oosterloo T., Fraternali F., Beccari G., 2013, A&A, 559, L11
Berenji B., Gaskins J., Meyer M., 2016, Physical Review D, 93, 045019
Berezhiani L., Khoury J., 2016, Physics Letters B, 753, 639
Bienaymé O., Famaey B., Wu X., Zhao H. S., Aubert D., 2009, A&A, 500, 801
Bílek M., Thies I., Kroupa P., Famaey B., 2017, A&A, accepted
Birkinshaw M., Gull S. F., 1983, Nature, 302, 315
Blaksley C., Bonamente M., 2010, New Astronomy, 15, 159
Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Gavazzi R., Moustakas L. A., Wayth R., Schlegel D. J., 2008, ApJ, 682, 964
Borsanyi S., et al., 2016, Nature, 539, 69
Borucki W. J., Koch D. G., Dunham E. W., Jenkins J. M., 1997, in Soderblom D., ed., Astronomical Society of the Pacific Conference Series Vol. 119, Planets Beyond the Solar System and the Next Generation of Space Missions. p. 153
Bournaud F., et al., 2007, Science, 316, 1166
Bovy J., Rix H.-W., 2013, ApJ, 779, 115
Bowden A., Evans N. W., Belokurov V., 2013, MNRAS, 435, 928
Brada R., Milgrom M., 1999, ApJ, 519, 590
Brimioulle F., Seitz S., Lerchster M., Bender R., Snigula J., 2013, MNRAS, 432, 1046
Broeils A. H., 1992, A&A, 256, 19
Brunthaler A., Reid M. J., Falcke H., Greenhill L. J., Henkel C., 2005, Science, 307, 1440
Cai Y.-C., Li B., Cole S., Frenk C. S., Neyrinck M., 2014, MNRAS, 439, 2978
Caldwell N., et al., 2017, ApJ, 839, 20
Candlish G. N., 2016, MNRAS, 460, 2571
Candlish G. N., Smith R., Fellhauer M., 2016, in Journal of Physics Conference Series. p. 012012, doi:10.1088/1742-6596/720/1/012012

Carignan C., Chemin L., Huchtmeier W. K., Lockman F. J., 2006, ApJL, 641, L109
Carlip S., 2001, Reports on Progress in Physics, 64, 885
Carr B., 1994, ARA&A, 32, 531
Casetti-Dinescu D. I., Girard T. M., Schriefer M., 2018, MNRAS, 473, 4064
Cautun M., Bose S., Frenk C. S., Guo Q., Han J., Hellwing W. A., Sawala T., Wang W., 2015, MNRAS, 452, 3838
Chae K.-H., Bernardi M., Sheth R. K., 2017, preprint, Arxiv (arXiv:1707.08280)
Chandrasekhar S., 1943, ApJ, 97, 255
Chesler P. M., Loeb A., 2017, Physical Review Letters, 119, 031102
Chiu M.-C., Ko C.-M., Tian Y., 2006, ApJ, 636, 565
Christian P., Loeb A., 2017, ApJL, 834, L20
Clark H. A., Lewis G. F., Scott P., 2016, MNRAS, 456, 1394
Clowe D., Gonzalez A., Markevitch M., 2004, ApJ, 604, 596
Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109
Collins M. L. M., et al., 2015, ApJL, 799, L13
Combes F., 2014, A&A, 571, A82
Cyburt R. H., Fields B. D., Olive K. A., Yeh T.-H., 2016, Rev. Mod. Phys., 88, 015004
Dalcanton J. J., et al., 2009, ApJS, 183, 67
Desmond H., 2017a, MNRAS, 464, 4160
Desmond H., 2017b, MNRAS, 472, L35
Duc P.-A., Paudel S., McDermid R. M., Cuillandre J.-C., Serra P., Bournaud F., Cappellari M., Emsellem E., 2014, MNRAS, 440, 1458
Duffy L. D., et al., 2006, Physical Review D, 74, 012006
Dyson F. W., Eddington A. S., Davidson C., 1920, Royal Society of London Philosophical Transactions Series A, 220, 291
Einstein A., 1905, Annalen der Physik, 322, 891
Einstein A., 1915, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847., pp 844–847
Faber S. M., Gallagher J. S., 1979, ARA&A, 17, 135
Famaey B., Binney J., 2005, MNRAS, 363, 603
Famaey B., McGaugh S. S., 2012, Living Reviews in Relativity, 15, 10
Famaey B., Bruneton J.-P., Zhao H., 2007, MNRAS, 377, L79
Famaey B., Khoury J., Penco R., 2018, JCAP, 3, 038
Fattahi A., et al., 2016, MNRAS, 457, 844
Fattahi A., Navarro J. F., Frenk C. S., Oman K., Sawala T., Schaller M., 2017, preprint, Arxiv (arXiv:1707.03898)
Fermi-LAT Collaboration 2015, Physical Review Letters, 115, 231301
Fernando N., Arias V., Guglielmo M., Lewis G. F., Ibata R. A., Power C., 2017, MNRAS, 465, 641
Fernando N., Arias V., Lewis G. F., Ibata R. A., Power C., 2018, MNRAS, 473, 2212
Fletcher R., Powell M. J. D., 1963, The Computer Journal, 6, 163
Flores H., Hammer F., Fouquet S., Puech M., Kroupa P., Yang Y., Pawlowski M., 2016, MNRAS, 457, L14
Francis C., Anderson E., 2014, Celestial Mechanics and Dynamical Astronomy, 118, 399
Freeman K. C., 1970, ApJ, 160, 811
Garaldi E., Romano-Díaz E., Borzyszkowski M., Porciani C., 2018, MNRAS, 473, 2234
Gentile G., Famaey B., Combes F., Kroupa P., Zhao H. S., Tiret O., 2007, A&A, 472, L25
Gentile G., Baes M., Famaey B., van Acoleyen K., 2010, MNRAS, 406, 2493
Gilmore G., Reid N., 1983, MNRAS, 202, 1025
Gómez P. L., et al., 2012, AJ, 144, 79
Gonzalez A. H., Clowe D., Bradač M., Zaritsky D., Jones C., Markevitch M., 2009, ApJ, 691, 525
Gregersen D., et al., 2015, AJ, 150, 189
Griest K., 1993, in Akerlof C. W., Srednicki M. A., eds, Annals of the New York Academy of Sciences Vol. 688, Texas/PASCOS '92: Relativistic Astrophysics and Particle Cosmology. p. 390 (arXiv:hep-ph/9303253), doi:10.1111/j.1749-6632.1993.tb43912.x
Hammer F., Yang Y. B., Flores H., Puech M., Fouquet S., 2015, ApJ, 813, 110
Hayden M. R., et al., 2015, ApJ, 808, 132
Hernandez X., Jiménez M. A., Allen C., 2012, European Physical Journal C, 72, 1884
Hodson A. O., Zhao H., 2017, A&A, 598, A127
Hohl F., 1971, ApJ, 168, 343
Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168
Ibata R. A., et al., 2013, Nature, 493, 62
Ibata N. G., Ibata R. A., Famaey B., Lewis G. F., 2014a, Nature, 511, 563
Ibata R. A., Ibata N. G., Lewis G. F., Martin N. F., Conn A., Elahi P., Arias V., Fernando N., 2014b, ApJL, 784, L6
Ibata R. A., Famaey B., Lewis G. F., Ibata N. G., Martin N., 2015, ApJ, 805, 67
Iocco F., Pato M., Bertone G., 2015, Physical Review D, 92, 084046
Iorio L., 2017, European Physical Journal C, 77, 149
Iršič V., et al., 2017, Physical Review D, 96, 023522
Jacobs B. A., Rizzi L., Tully R. B., Shaya E. J., Makarov D. I., Makarova L., 2009, AJ, 138, 332
Jansen F., et al., 2001, A&A, 365, L1
Jayaraman A., Gilmore G., Wyse R. F. G., Norris J. E., Belokurov V., 2013, MNRAS, 431, 930
Jee M. J., Hughes J. P., Menanteau F., Sifón C., Mandelbaum R., Barrientos L. F., Infante L., Ng K. Y., 2014, ApJ, 785, 20
Jungman G., Kamionkowski M., Griest K., 1996, Physics Reports, 267, 195
Jurić M., et al., 2008, ApJ, 673, 864
Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2012, ApJ, 761, 98
Kahn F. D., Woltjer L., 1959, ApJ, 130, 705
Kamada A., Kaplinghat M., Pace A. B., Yu H.-B., 2017, Physical Review Letters, 119, 111102
Kamionkowski M., 1998, in Gava E., Masiero A., Narain K. S., Randjbar-Daemi S., Senjanovic G., Smirnov A., Shafi Q., eds, High Energy Physics and Cosmology, 1997 Summer School. p. 394 (arXiv:hep-ph/9710467)
Katz H., McGaugh S., Teuben P., Angus G. W., 2013, ApJ, 772, 10
Keller B. W., Wadsley J. W., 2017, ApJL, 835, L17
Keller B. W., Wadsley J., Couchman H. M. P., 2016, MNRAS, 463, 1431
Kerins E. J., Carr B. J., 1995, Nuclear Physics B: Proceedings Supplements, 43, 157
Kervella P., Mignard F., Mérand A., Thévenin F., 2016, A&A, 594, A107
Kervella P., Thévenin F., Lovis C., 2017, A&A, 598, L7
Khoury J., 2016, Physical Review D, 93, 103533
Kirby E. N., Bullock J. S., Boylan-Kolchin M., Kaplinghat M., Cohen J. G., 2014, MNRAS, 439, 1015
Klimentowski J., Łokas E. L., Knebe A., Gottlöber S., Martinez-Vaquero L. A., Yepes G., Hoffman Y., 2010, MNRAS, 402, 1899
Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Komatsu E., et al., 2011, ApJS, 192, 18
Kormendy J., Drory N., Bender R., Cornell M. E., 2010, ApJ, 723, 54
Kraljic D., Sarkar S., 2015, JCAP, 4, 050
Kroupa P., 2015, Canadian Journal of Physics, 93, 169
Kroupa P., Theis C., Boily C. M., 2005, A&A, 431, 517

Kunder A., et al., 2017, AJ, 153, 75
LIGO Collaboration 2016, Physical Review Letters, 116, 061102
LUX Collaboration 2017, Physical Review Letters, 118, 021303
Lage C., Farrar G., 2014, ApJ, 787, 144
Lelli F., McGaugh S. S., Schombert J. M., 2016, AJ, 152, 157
Lelli F., McGaugh S. S., Schombert J. M., 2017a, MNRAS, 468, L68
Lelli F., McGaugh S. S., Schombert J. M., Pawlowski M. S., 2017b, ApJ, 836, 152
Lena D., Robinson A., Marconi A., Axon D. J., Capetti A., Merritt D., Batcheldor D., 2014, ApJ, 795, 146
Lenoir B., Christophe B., Reynaud S., 2011, preprint, Arxiv (arXiv:1107.0861)
Leonard P. J. T., Tremaine S., 1990, ApJ, 353, 486
Lin C. C., Shu F. H., 1964, ApJ, 140, 646
López-Corredoira M., Kroupa P., 2016, ApJ, 817, 75
Lüghausen F., Famaey B., Kroupa P., 2015, Canadian Journal of Physics, 93, 232
Ma C., et al., 1998, AJ, 116, 516
Ma J., Wu Z., Wang S., Fan Z., Zhou X., Wu J., Jiang Z., Chen J., 2010, PASP, 122, 1164
Maji M., Zhu Q., Marinacci F., Li Y., 2017, preprint, Arxiv (arXiv:1702.00497)
Makarov D. I., Makarova L. N., Uklein R. I., 2013, Astrophysical Bulletin, 68, 125
Martig M., Bournaud F., Croton D. J., Dekel A., Teyssier R., 2012, ApJ, 756, 26
Massey P., Henning P. A., Kraan-Korteweg R. C., 2003, AJ, 126, 2362
Mastropietro C., Burkert A., 2008, MNRAS, 389, 967
McConnachie A. W., 2012, AJ, 144, 4
McConnachie A. W., et al., 2009, Nature, 461, 66
McGaugh S. S., 1996, MNRAS, 280, 337
McGaugh S. S., 2008, ApJ, 683, 137
McGaugh S. S., 2011, Physical Review Letters, 106, 121303
McGaugh S. S., 2016a, ApJ, 816, 42
McGaugh S. S., 2016b, ApJL, 832, L8
McGaugh S., Milgrom M., 2013, ApJ, 775, 139
McGaugh S. S., Wolf J., 2010, ApJ, 722, 248
McGaugh S. S., Schombert J. M., Bothun G. D., 1995, AJ, 109, 2019
McGaugh S., Lelli F., Schombert J., 2016, Phys. Rev. Lett., 117, 201101
McMillan P. J., 2011, MNRAS, 414, 2446
McMillan P. J., 2017, MNRAS, 465, 76
McQuinn K. B. W., et al., 2015, ApJ, 812, 158
Menanteau F., et al., 2012, ApJ, 748, 7
Merali Z., 2016, Science, 352, 1040
Milgrom M., 1983, ApJ, 270, 365
Milgrom M., 1986, ApJ, 302, 617
Milgrom M., 1999, Phys. Lett. A, 253, 273
Milgrom M., 2010, MNRAS, 403, 886
Milgrom M., 2013, Physical Review Letters, 111, 041105
Milgrom M., 2014, Physical Review D, 89, 024027
Milgrom M., 2016, preprint, Arxiv (arXiv:1610.07538)
Mirabel I. F., Dottori H., Lutz D., 1992, A&A, 256, L19
Molnar S. M., Broadhurst T., 2015, ApJ, 800, 37
Molnar S. M., Chiu I.-N. T., Broadhurst T., Stadel J. G., 2013a, ApJ, 779, 63
Molnar S. M., Chiu I.-N. T., Broadhurst T., Stadel J. G., 2013b, ApJ, 779, 63
Müller O., Pawlowski M. S., Jerjen H., Lelli F., 2018, Science, 359, 534
NANOGrav Collaboration 2018, preprint, Arxiv (arXiv:1801.02617)
Nesseris S., Pantazis G., Perivolaropoulos L., 2017, Physical Review D, 96, 023542
Nicastro F., Senatore F., Krongold Y., Mathur S., Elvis M., 2016, ApJL, 828, L12
Noguchi M., 1999, ApJ, 514, 77
Norris M. A., et al., 2016, ApJ, 832, 198
Okazaki T., Taniguchi Y., 2000, ApJ, 543, 149
Olling R. P., Merrifield M. R., 2001, MNRAS, 326, 164
Ostriker J. P., Peebles P. J. E., 1973, ApJ, 186, 467
Ostriker J. P., Steinhardt P. J., 1995, Nature, 377, 600
Paczynski B., 1986, ApJ, 304, 1
PandaX-II Collaboration 2016, Phys. Rev. Lett., 117, 121303
Papastergis E., Adams E. A. K., van der Hulst J. M., 2016, A&A, 593, A39
Pawlowski M. S., 2016, MNRAS, 456, 448
Pawlowski M. S., Kroupa P., 2013, MNRAS, 435, 2116
Pawlowski M. S., et al., 2014, MNRAS, 442, 2362
Pawlowski M. S., Famaey B., Merritt D., Kroupa P., 2015, ApJ, 815, 19
Pawlowski M. S., et al., 2017, Astronomische Nachrichten, 338, 854
Pazy E., 2013, Phys. Rev. D, 87, 084063
Peñarrubia J., Fattahi A., 2017, MNRAS, 468, 1300
Peñarrubia J., Ma Y.-Z., Walker M. G., McConnachie A., 2014, MNRAS, 443, 2204
Peebles P. J. E., 2017, preprint, Arxiv (arXiv:1705.10683)
Peebles P. J. E., Tully R. B., 2013, preprint, Arxiv (arXiv:1302.6982)
Peebles P. J. E., Tully R. B., Shaya E. J., 2011, preprint, Arxiv (arXiv:1105.5596)
Peng Y., Maiolino R., Cochrane R., 2015, Nature, 521, 192
Perryman M. A. C., et al., 2001, A&A, 369, 339
Phelps S., Nusser A., Desjacques V., 2013, ApJ, 775, 102
Piffl T., et al., 2014, A&A, 562, A91
Pittordis C., Sutherland W., 2017, preprint, Arxiv (arXiv:1711.10867)
Planck Collaboration XIII 2016, A&A, 594, A13
Planck Collaboration XVI 2014, A&A, 571, A16
Planck Collaboration XXVII 2014, A&A, 571, A27
Plummer H. C., 1911, MNRAS, 71, 460
Privon G. C., Barnes J. E., Evans A. S., Hibbard J. E., Yun M. S., Mazzarella J. M., Armus L., Surace J., 2013, ApJ, 771, 120
Quillen A. C., Garnett D. R., 2001, in Funes J. G., Corsini E. M., eds, Astronomical Society of the Pacific Conference Series Vol. 230, Galaxy Disks and Disk Galaxies. pp 87–88 (arXiv:astro-ph/0004210)
Refsdal S., 1966, MNRAS, 134, 315
Riess A. G., et al., 1998, AJ, 116, 1009
Roberts M. S., Whitehurst R. N., 1975, ApJ, 201, 327
Rogstad D. H., Shostak G. S., 1972, ApJ, 176, 315
Rubin V. C., Ford Jr. W. K., 1970, ApJ, 159, 379
Ruskin S., 2017, American Journal of Physics, 85, 159
Ryu T., Perna R., Haiman Z., Ostriker J. P., Stone N. C., 2018, MNRAS, 473, 3410
Salem M., Besla G., Bryan G., Putman M., van der Marel R. P., Tonnesen S., 2015, ApJ, 815, 77
Sales L. V., Navarro J. F., Abadi M. G., Steinmetz M., 2007, MNRAS, 379, 1475
Salucci P., Turini N., 2017, preprint, Arxiv (arXiv:1707.01059)
Salucci P., Lapi A., Tonini C., Gentile G., Yegorova I., Klein U., 2007, MNRAS, 378, 41
Sancisi R., 2004, in Ryder S., Pisano D., Walker M., Freeman K., eds, IAU Symposium Vol. 220, Dark Matter in Galaxies. p. 233 (arXiv:astro-ph/0311348)
Sandage A., 1986, ApJ, 307, 1
Sanders R. H., 2003, MNRAS, 342, 901
Sarazin C. L., 1986, Reviews of Modern Physics, 58, 1
Scarpa R., Ottolina R., Falomo R., Treves A., 2017, International Journal of Modern Physics D, 26, 1750067
Schechter P., 1976, ApJ, 203, 297

Schmidt K. H., 1958, Astronomische Nachrichten, 284, 76
Schönrich R., 2012, MNRAS, 427, 274
Shannon R. M., et al., 2015, Science, 349, 1522
Shappee B. J., Stanek K. Z., 2011, ApJ, 733, 124
Sikivie P., 1983, Physical Review Letters, 51, 1415
Slipher V. M., 1913, Lowell Observatory Bulletin, 2, 56
Smolin L., 2017, Physical Review D, 96, 083523
Snaith O. N., Haywood M., Di Matteo P., Lehnert M. D., Combes F., Katz D., Gómez A., 2014, ApJL, 781, L31
Sohn S., 2016, Proper Motions of the Crater-Leo Group: Testing the Group Infall Scenario, HST Proposal
Sohn S. T., et al., 2017, ApJ, 849, 93
Steigman G., Turner M. S., 1985, Nuclear Physics B, 253, 375
Swaters R. A., Sancisi R., van Albada T. S., van der Hulst J. M., 2009, A&A, 493, 871
Tenneti A., Mao Y.-Y., Croft R. A. C., Di Matteo T., Kosowsky A., Zago F., Zentner A. R., 2018, MNRAS, 474, 3125
Teyssier R., 2002, A&A, 385, 337
Theia Collaboration 2017, preprint, Arxiv (arXiv:1707.01348)
Thompson R., Nagamine K., 2012, MNRAS, 419, 3560
Tian Y., Ko C.-M., 2017, MNRAS, 472, 765
Tisserand P., et al., 2007, A&A, 469, 387
Torrealba G., Koposov S. E., Belokurov V., Irwin M., 2016, MNRAS, 459, 2370
Tucker W. H., Tananbaum H., Remillard R. A., 1995, ApJ, 444, 532
Tucker W., et al., 1998, ApJL, 496, L5
Tully R. B., 2015, AJ, 149, 54
Tully R. B., Fisher J. R., 1977, A&A, 54, 661
Tully R. B., et al., 2013, AJ, 146, 86
Turyshev S. G., Toth V. T., Kinsella G., Lee S.-C., Lok S. M., Ellis J., 2012, Physical Review Letters, 108, 241101
Verlinde E. P., 2017, SciPost Physics, 2, 016
Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, Physical Review D, 88, 043502
Virgo & LIGO Collaborations 2017, Physical Review Letters, 119, 161101
Wang Y., Mohanty S. D., 2017, Physical Review Letters, 118, 151104
Wang L., Greene J. E., Ju W., Rafikov R. R., Ruan J. J., Schneider D. P., 2017, ApJ, 834, 129
Wetzstein M., Naab T., Burkert A., 2007, MNRAS, 375, 805
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Williams A. A., Belokurov V., Casey A. R., Evans N. W., 2017a, MNRAS, 468, 2359
Williams B. F., et al., 2017b, ApJ, 846, 145
Wu X., Famaey B., Gentile G., Perets H., Zhao H., 2008, MNRAS, 386, 2199
York D. G., et al., 2000, AJ, 120, 1579
Zacharias N., Finch C., Frouard J., 2017, AJ, 153, 166
Zanella A., et al., 2015, Nature, 521, 54
Zhao H., Famaey B., 2012, Physical Review D, 86, 067301
Zhao H., Famaey B., Lüghausen F., Kroupa P., 2013, A&A, 557, L3
Zwicky F., 1937, ApJ, 86, 217
van den Bergh S., 1999, ApJL, 517, L97
van der Marel R. P., Kallivayalil N., 2014, ApJ, 781, 121
van der Marel R. P., Fardal M., Besla G., Beaton R. L., Sohn S. T., Anderson J., Brown T., Guhathakurta P., 2012a, ApJ, 753, 8
van der Marel R. P., Besla G., Cox T. J., Sohn S. T., Anderson J., 2012b, ApJ, 753, 9

This thesis has been typeset from a TEX/LATEX file prepared by the author.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **450,** 3155–3168 (2015) doi:10.1093/mnras/stv802

Effects of lens motion and uneven magnification on image spectra

Indranil Banik★ and Hongsheng Zhao★

Scottish Universities Physics Alliance, University of St Andrews, North Haugh, St Andrews, Fife KY16 9SS, UK

Accepted 2015 April 9. Received 2015 April 3; in original form 2014 December 15

ABSTRACT

Counter to intuition, the images of an extended galaxy lensed by a moving galaxy cluster should have slightly different spectra in any metric gravity theory. This is mainly for two reasons. One relies on the gravitational potential of a moving lens being time dependent (the moving cluster effect, MCE). The other is due to uneven magnification across the extended, rotating source (the differential magnification effect, DME). The time delay between the images can also cause their redshifts to differ because of cosmological expansion. This differential expansion effect is likely to be small. Using a simple model, we derive these effects from first principles. One application would be to the Bullet Cluster, whose large tangential velocity may be inconsistent with the Λ cold dark matter paradigm. This velocity can be estimated with complicated hydrodynamic models. Uncertainties with such models can be avoided using the MCE. We argue that the MCE should be observable with Atacama Large Millimetre Array. However, such measurements can be corrupted by the DME if typical spiral galaxies are used as sources. Fortunately, we find that if detailed spectral line profiles were available, then the DME and MCE could be distinguished. It might also be feasible to calculate how much the DME should affect the mean redshift of each image. Resolved observations of the source would be required to do this accurately. The DME is of order the source angular size divided by the Einstein radius times the redshift variation across the source. Thus, it mostly affects nearly edge-on spiral galaxies in certain orientations. This suggests that observers should reduce the DME by careful choice of target, a possibility we discuss in some detail.

Key words: gravitational lensing: strong – galaxies: clusters: individual: 1E 0657−56 – galaxies: kinematics and dynamics – dark matter.

1 INTRODUCTION

The standard Λ cold dark matter (ΛCDM) paradigm (Ostriker & Steinhardt 1995) still faces many challenges in reproducing galaxy scale observations (for a recent review, see Famaey & McGaugh 2012). Particularly problematic is the anisotropic distribution of satellites around Local Group galaxies, a question recently revisited in detail (Pawlowski et al. 2014). A different analysis focusing on Andromeda came to similar conclusions (Ibata et al. 2014). The relevant observations for the Milky Way (Pawlowski & Kroupa 2013) and Andromeda (Ibata et al. 2013) are difficult to repeat outside the Local Group because of the need to obtain 3D positions and velocities.

On a larger scale, Cai et al. (2014) found that the collision speed distribution of interacting galaxy clusters can be quite sensitive to the underlying law of gravitation. Thus, the high collision speed of the components of the Bullet Cluster 1E 0657−56 (Tucker, Tananbaum & Remillard 1995) has been argued in favour of modified gravity (Katz et al. 2013). However, this speed is not directly measured as the collision is mostly in the plane of the sky. Instead, the speed is estimated using simulations of the shock generated in the gas by the collision. The separation of the dark matter (DM) and gas (Clowe et al. 2006) also plays an important role – there is less gas drag at lower speeds, so the separation is generally reduced.

A collision speed close to 3000 km s^{-1} is thought to be required to explain the observed properties of the Bullet Cluster (Mastropietro & Burkert 2008). For the inferred masses of the components (Clowe, Gonzalez & Markevitch 2004), this appears difficult to reconcile with ΛCDM (Thompson & Nagamine 2012). This work suggested that a cosmological simulation requires a comoving volume of $(4.48\,h^{-1}\ \mathrm{Gpc})^3$ to see an analogue to the Bullet Cluster.

The recent work of Lage & Farrar (2014) finds a similar collision speed but suggests a higher mass for the Bullet Cluster's components. While higher mass objects are likely to collide faster, such heavy clusters are rare in cosmological simulations. For example, their own unpublished work (Lage & Farrar 2015) based on the Horizon Run *N*-body simulation (Kim et al. 2009) showed how there were only seven cluster pairs with masses comparable to

★ E-mail: ib45@st-andrews.ac.uk (IB); hz4@st-andrews.ac.uk (HZ)

their higher estimate for the Bullet Cluster mass. Because a larger volume of $(6.59\,h^{-1}\,\mathrm{Gpc})^3$ was used in the simulation and $\left(\frac{6.59}{4.48}\right)^3 \approx 3$, this result is not very surprising in light of previous works.

However, some recent unpublished studies have raised the probability estimate of observing a galaxy cluster merger with properties comparable to the Bullet Cluster. Bouillot et al. (2014) used a larger box size of $(21\,h^{-1}\,\mathrm{Gpc})^3$, using the Dark Energy Universe Simulation-Full Universe Run (DEUS-FUR) simulation. Thompson, Davé & Nagamine (2014) took issue with the friends-of-friends algorithm (Davis et al. 1985) long used to search outputs of *N*-body simulations for analogues to the Bullet Cluster. After switching to the recently developed ROCKSTAR algorithm (Behroozi, Wechsler & Wu 2013), the rate of occurrence of analogues to the Bullet Cluster increased by a factor of ∼100. Despite this, Thompson et al. (2014) quoted the probability of a collision as fast as the observed one as only one in 2170, which is still fairly small.

Moreover, a few other massive colliding clusters with high infall velocities have been discovered in the last few years (Gómez et al. 2012; Menanteau et al. 2012; Molnar et al. 2013b). The El Gordo Cluster (ACT-CL J0102−4915) may be particularly problematic due to its combination of high redshift ($z = 0.87$; Menanteau et al. 2012), high mass (Jee et al. 2014) and high inferred collision speed (Molnar & Broadhurst 2015).

A detailed analysis of how likely it is that observers would have seen interacting clusters with the observed properties is still lacking. One would need to account for incomplete sky coverage and perhaps faster collisions being easier to discover due to greater shock heating of the gas. A key input to any such analysis must be the collision speeds of the components. This work focuses on measuring cluster motions more accurately.

Molnar et al. (2013b) argue that inferring collision speeds from observations of the shock can be non-trivial just due to projection effects, let alone other complexities of baryonic physics. To see if there is any tension with the ΛCDM model, the collision speeds should be determined in a more direct way. Ultimately, we would like to determine the proper motions of colliding clusters.

Although not feasible by traditional methods, such motions may be inferred using the moving cluster effect (MCE; Birkinshaw & Gull 1983). As derived later, this effect relies on the gravitational potential of an object being time dependent due to its motion. Consequently, if a source behind the object were multiply imaged, the images would have slightly different redshifts. Moreover, as the DM generally outweighs the gas on cluster scales (Blaksley & Bonamente 2010), the MCE is mostly sensitive to the motion of the DM. This is simpler to model than the gas, making the results easier to compare with simulations.

The MCE would likely be around 1 km s^{-1} for the Bullet Cluster (Molnar et al. 2013a). The effect may be searched for using cosmic microwave background (CMB) photons (e.g. Cai et al. 2010). However, as noted by those authors, temperature anisotropies in the CMB make it difficult to spot such a signal around an individual object. Thus, we focus instead on using a multiply imaged background galaxy as the source. Spectral features in this galaxy could be used to determine the redshifts of its multiple images.

We consider the feasibility of obtaining measurements of the required accuracy using Atacama Large Millimetre Array (ALMA) in Section 7. Measuring this effect seems to be within our reach. One might instead conduct the observations in the visible/near-infrared (IR) with large spectroscopic instruments such as the Thirty Meter Telescope (TMT) and European-Extremely Large Telescope (E-ELT).

Therefore, it is important to consider other effects that might also cause the redshifts of double images to differ. Perhaps the most important such mechanism is what we term the differential magnification effect (DME). This depends on details of the source. If this is a rotating disc galaxy not viewed face-on, then different parts of the source have different radial velocities and hence redshifts.

The lens magnifies the source non-uniformly. The exact way in which this occurs is different for each image. Consequently, the intensity-weighted mean redshift of the images is usually different.

If one could perform integral field spectroscopy of the source galaxy accurate to ∼1 km s^{-1}, then one would simply need to compare the redshift of the same part of the galaxy between the two images. By focusing on a small part of the galaxy, the DME would be greatly reduced. However, this will be a challenging observational goal. The high spectral accuracy demanded by MCE measurements means the source will likely be spatially unresolved in the near future.

Assuming this to be the case, we determine the order of magnitude of the DME for a typical disc galaxy. We find that it may well be significant. Thus, we explore exactly how it affects the profiles of individual spectral lines. The effect is quite different to the MCE, which simply shifts each line. This may provide a way to correct for the DME and also to verify that a redshift difference is indeed caused by the MCE. Without spectra detailed enough to see such small differences between line profiles, it might still be possible to calculate the DME, though the determination would be less secure.

The additional complications and uncertainty introduced by trying to correct for the DME necessitate a discussion on how it may be reduced. Aside from the obvious steps of using ellipticals/face-on spirals and smaller – likely slower rotating – galaxies, an important factor to consider is how much the magnification varies across the source. The larger the variation, the larger the DME.

For this reason, an edge-on fast-rotating spiral galaxy might still be a good target if it is oriented so the magnification is nearly constant across the image. At the other extreme, the magnification varies rapidly near a caustic. Therefore, caustic images are likely to be strongly affected by the DME (Molnar et al. 2013a).

We emphasize the need to model both the redshift structure of the source and the deflection map of the lens when trying to use precise lensed image redshifts to determine the tangential motion of the lens.

2 THE MOVING CLUSTER EFFECT IN STANDARD AND NON-STANDARD GRAVITY

2.1 The lensing geometry

Fig. 1 illustrates the basic geometry that will be considered here. Because we are mostly concerned with angles on the sky, the distances relevant to us are the angular diameter distances to the lens and source (D_l and D_s). Also important is the angular diameter distance to the source galaxy as perceived by an astronomer at the lens, measured at the epoch that the lens is currently observed at (z_l). This last distance we denote D_ls.

2.2 Including lens motion

Suppose that the lens moves transversely to $\boldsymbol{OL}$. Thus, one of the light paths gets 'stretched' while the other gets 'squeezed', leading to a redshift difference between the images. To calculate this effect, we make the thin lens or triangle approximation whereby each photon trajectory is treated as two straight lines. In this case, the

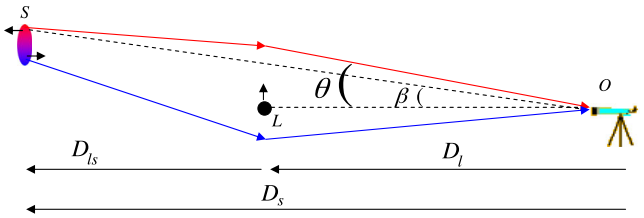


Figure 1. The lensing geometry is depicted here. Upper photon trajectory = primary image (same side as unlensed source), lower trajectory = secondary image. Relevant distances are indicated at bottom. The lens is treated as a point mass moving transversely to the viewing direction. The source is an extended disc galaxy. There is a redshift gradient across it due to rotation.

light arrival time surface is generally given by (Kovner 1990)

$$cT(\boldsymbol{\theta}) = \text{constant} + \frac{D_l D_s}{2D_{ls}}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \Psi(D_l\boldsymbol{\theta} - \boldsymbol{x}_1). \quad (1)$$

This consists of a geometric part and a relativistic part due to the lensing potential Ψ. We assume that Ψ depends only on position relative to the lens, which is located at $\boldsymbol{x}_1$ relative to some reference point in the lens plane.

The path length $cT(\boldsymbol{\theta})$ can be thought of as a function of the lens plane position $\boldsymbol{\theta}$ hit by a ray from the source. The actual rays for the images are at the extrema of this function (Fermat's principle).

We briefly describe how the geometric part of equation (1) is derived. Each section of a hypothetical undeviated photon trajectory can be mapped on to a part of the actual trajectory. The latter is slightly longer as there is an extra factor of the secant of the angle between them. This is expanded at second order, as the angle is small. The angle is $(\theta - \beta)$ for the stretch LO while for OS, it is $(\theta - \beta)\,\frac{D_l}{D_{ls}}\,\frac{1+z_l}{1+z_s}$. The last factor arises because photons emitted in the same direction gradually get further apart due to cosmic expansion. Thus, photons emitted in different directions end up more widely separated than in a static universe.

Equation (1) follows most naturally if using comoving distances, which can be added simply. This fact leads to the very useful relation between angular diameter distances:

$$(1 + z_l)\, D_l + (1 + z_s)\, D_{ls} = (1 + z_s)\, D_s. \quad (2)$$

If the lens moves in the transverse direction, then $\boldsymbol{x}_1$ changes and so the lensing potential Ψ changes at every point in the lens plane. Thus, transverse motion of the lens would cause the arrival time to change according to

$$c\dot{T} = \dot{\boldsymbol{x}}_1 \cdot \nabla\Psi \equiv -\boldsymbol{v}_t \cdot \boldsymbol{\alpha}. \quad (3)$$

$\boldsymbol{\alpha} \equiv -\nabla\Psi$ is the unreduced (true) deflection angle and $\boldsymbol{v}_t \equiv \dot{\boldsymbol{x}}_1$ is the transverse velocity of the lens (note time here refers to that measured by a clock at the lens).

The rate of increase of the path length $c\dot{T}$ is equivalent to a shift of the intrinsic spectrum of the source. While the source's intrinsic spectrum cannot be directly measured, the relative spectra of the images in a multiple image system can. Images 1 and 2 would have a relative redshift velocity:

$$\delta V_r \equiv V_1 - V_2 = -\boldsymbol{v}_t \cdot (\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2), \quad (4)$$

where $\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2$ is the relative deflection angle between the images. The observed angular separation between the images is $\frac{D_{ls}}{D_s}$ times as much. This allows the MCE to be calculated without knowing what the deflection angles are, as long as one is sure of the identification of the double image and the distances to the lens and source.

So far, we have used time measured by a clock at the lens. Using one on the Earth instead would introduce a factor of $(1 + z_l)$ to the time delay. Putting it in, we should think of $\boldsymbol{v}_t$ as the transverse peculiar velocity in comoving coordinates. This is the comoving lens distance times the relative proper motion of the lens with respect to that of the source.[1] This takes account of transverse motions of the observer and the source. In Section 3.2, we show that such motions affect image redshifts much less than motion of the lens.

For multiple lens planes, one would simply add the redshift differences due to each plane.

The above derivation is a property of metric theories of gravity. Hence, it is independent of details of the theory, something we now show. Consider a static Universe with no observer–source relative motion. Use a reference frame moving with the lens, so the source and observer both appear to move at $-\dot{\boldsymbol{x}}_1$. In general, there is a Doppler shift for a photon emitted by the source as perceived at the lens. A similar effect arises between lens and observer. The shifts cancel if the photon is not deflected by the lens (if the source emits a photon 'backward', then the observer 'ploughs into' the photon).

However, if the photon is deflected, then there is a net frequency shift between source and observer. This shift is different for another photon which gets deflected by a different amount. Therefore, the difference in deflection angles determines the redshift difference between the photons. If observers could be sure the photons initially had the same energy, then the redshift difference could be directly measured, thus constraining $\dot{\boldsymbol{x}}_1$.

For small deflections by a non-relativistic lens, the result in an expanding Universe is exactly the same as in a static one, once all the angles have been properly accounted for.

2.3 Point mass lenses

The unlensed source and images all lie along a line, so we only consider positions along this line. For a point mass lens, one can generically say that a ray of light with impact parameter $D_l\theta$ is deflected by

$$\alpha = -\frac{4G\widetilde{M}}{c^2 D_l \theta}, \quad (5)$$

where $\widetilde{M}$ is the 'equivalent lensing mass' in alternative gravity. Different gravity theories require different amount of real mass M to produce the observed equivalent mass $\widetilde{M}$, the Einstein radius θ_E or the deflection angles $\alpha_{1,2}$. In general, $\widetilde{M}$ depends on position in modified gravity theories, even with a point mass. For simplicity, we neglect this.

Combining equations (4) and (5), we get that

$$V_1 - V_2 = \frac{4G\widetilde{M} v_t}{c^2 D_l}\left(\frac{1}{\theta_1} - \frac{1}{\theta_2}\right). \quad (6)$$

Note that the signs of θ_1 and θ_2 are opposite because the images are on opposite sides of the lens (equation 10).

v_t is the component of the transverse velocity $\boldsymbol{v}_t$ projected along the line connecting images 1 and 2. If the lens proper motion is orthogonal to the image separation, then v_t would be zero.

[1] Physically, it would be the peculiar velocity of the lens in the direction orthogonal to our line of sight, in the absence of peculiar motions of either observer or source.

Noting that deflecting a photon at the lens only affects part of its trajectory, we get the classical lens equation

$$\beta = \theta - \frac{D_{\rm ls}}{D_{\rm s}} \frac{4G\widetilde{M}}{c^2 D_{\rm l} \theta} \tag{7}$$

$$\equiv \theta - \frac{{\theta_{\rm E}}^2}{\theta}, \quad \text{where } \theta_{\rm E} \equiv \sqrt{\frac{4G\widetilde{M}}{c^2} \frac{D_{\rm ls}}{D_{\rm l} D_{\rm s}}}. \tag{8}$$

The Einstein radius $\theta_{\rm E}$ defines a typical angular scale for the problem. It will be convenient to use this to normalize all relevant angles. Thus, we let

$$u \equiv \frac{\beta}{\theta_{\rm E}} \quad \text{and} \quad y \equiv \frac{\theta}{\theta_{\rm E}}. \tag{9}$$

Images are formed where

$$y = \frac{1}{2} \left(u \pm \sqrt{u^2 + 4} \right). \tag{10}$$

For later use, we note that the magnification of a small part of the source located at an unlensed angular position of $\beta \equiv u\theta_{\rm E}$ relative to the lens is given by

$$A = \left| \frac{\theta}{\beta} \frac{\partial \theta}{\partial \beta} \right| \tag{11}$$

$$= \frac{1}{2} \left(\frac{u^2 + 2}{u\sqrt{u^2 + 4}} \pm 1 \right). \tag{12}$$

The result follows from the surface brightness of an unlensed source and a lensed one being equal (Liouville's theorem). Thus, A is the Jacobian of the mapping between where objects appear on the sky and where they would without a lens. The modulus signs are needed because otherwise $A < 0$ for the secondary image (indicating that it is inverted).

The secondary image becomes very faint if $u \gg 1$ (in this case, $A \sim \frac{2}{u^4}$). It is difficult to find a source with $u \ll 1$ as this corresponds to a very small part of the sky. Thus, any source used for MCE measurements will very likely have $u \sim 1$. We assume this is the case.

3 CAUSES OF DIFFERENCES BETWEEN MULTIPLE IMAGE REDSHIFTS

Before deriving the DME, we briefly consider a few factors that can affect redshift differences between double images of a source strongly lensed by the Bullet Cluster.

3.1 Lens motion

The MCE is maximal for a source displaced from the lens in the direction of its proper motion. This direction can be determined from images of the shock fairly easily. Moreover, observers should select targets to maximize the MCE. Thus, we assume the double images are indeed separated along the direction of motion of the lens. Otherwise, the MCE is reduced by the cosine of the angle between them (equation 4).

With these assumptions, we combine equations (6) and (10) to get that the difference in redshift velocity between the two images is

$$\Delta \overline{v_{\rm r}}|_{\rm MCE} = \frac{2v_{\rm t}\sqrt{GM\left(u^2+4\right)}}{c} \sqrt{\frac{D_{\rm s}}{D_{\rm ls} D_{\rm l}}}. \tag{13}$$

Table 1. Input parameters used for Fig. 4. The source galaxy is assumed positioned so as to maximize the MCE (i.e. it is separated from the lens on the sky along the direction of motion of the lens, which is clear from images). The lens mass should roughly correspond to the subcluster in the Bullet. A flat ΛCDM cosmology is adopted (Planck Collaboration XVI 2014).

Parameter	Meaning	Value
H_0	Present Hubble constant	67.3 km s^{-1} Mpc^{-1}
$\Omega_{\rm m}$	Present matter density	0.315
$D_{\rm l}$	(Angular diameter) Distance to lens at $z_{\rm l} = 0.296$	0.945 Gpc
$D_{\rm s}$	Distance to source at $z_{\rm s} = 1.7$	1.795 Gpc
$D_{\rm ls}$	Distance to source from lens position in space–time	1.341 Gpc
M	Mass of lens	1.2×10^{14} M$_\odot$
$r_{\rm d}$	Scale length of source galaxy	3.068 kpc
$v_{\rm t}$	Tangential velocity of lens	3000 km s^{-1}
$v_{\rm f}$	Flat line level of source galaxy rotation curve	100 km s^{-1}
$\sin i \cos\gamma$	See Fig. 2 and equation (21)	$\frac{1}{2}$

Using parameters appropriate to the Bullet Cluster (Table 1), the effect is around 1 km s^{-1}.

3.2 Source and observer motion

The peculiar velocity (with respect to the CMB) of the Sun is well known (369 km s^{-1}; Planck Collaboration XXVII 2014) and could be included in a more careful analysis. We do not include it as we only seek a rough idea of the magnitude of the MCE. This will not be substantially affected by observer motion as this is much slower than that of the lens (∼3000 km s^{-1}).

More problematic may be the unknown peculiar velocity of the source. Treating the Local Group peculiar velocity (∼630 km s^{-1}) as typical for galaxies, the Bullet Cluster transverse motion likely greatly exceeds the source's peculiar velocity. In this case, only the component of this velocity parallel to the lens transverse motion has much effect on image redshifts, leading to a factor of $\frac{1}{2}$ on average.[2]

Another factor of $\frac{1+z_{\rm l}}{1+z_{\rm s}} \frac{D_{\rm l}}{D_{\rm s}} \approx \frac{1}{4}$ arises due to the geometry of the situation and cosmic expansion. Moreover, typical peculiar velocities were smaller long ago. Supposing they were $\frac{1}{2}$ as much as today at $z_{\rm s} = 1.5$, we see that source motion cannot affect the inferred lens velocity much more than ∼50 km s^{-1}. This effect can be reduced by observing more than one double image pair. However, we consider the accuracy with even just one well-observed pair sufficient.

Thus, we ignore any motion other than that of the lens. We note that it might be good to avoid source galaxies which are interacting, as their peculiar velocities might be higher.

3.3 Cosmological expansion

A redshift difference between the images can also arise because the time of flight of photons emitted by the source is different depending on which path they took to get to the Earth. As the photons for both images arrive simultaneously, the photons for one image must have been emitted earlier than for the other. Thus, in an expanding Universe, one of the images will have a higher redshift. Because of both a longer path length and a stronger gravitational

[2] For an angle between a fixed vector and another statistically isotropic one, $\langle|\cos\theta|\rangle = \frac{1}{2}$.

field along the path, this image is the secondary (on the opposite side of the lens as the unlensed source would appear – see Fig. 1). We term this the differential expansion effect (DEE).

A quick way to estimate the DEE is to assume that cosmological distances like $D_{\rm L}$ are usually $\sim\frac{c}{H}$. The extra path length $\sim D_{\rm L}\theta^2$. The effect of the difference in gravitational time delays can crudely be approximated as equal to that due to different geometric path lengths.

The DEE expressed as a redshift $= H\Delta t \approx \theta^2$. Meanwhile, the MCE $\sim\frac{v}{c}\theta$. Assuming a velocity of 3000 km s^{-1} and an image separation of 20 arcsec, we see that the MCE is $\sim$50 times larger than the DEE. Thus, we calculate the DEE more precisely.

We first consider just the difference in time of flight ('delay') due to different strengths of gravity along the two possible photon paths (Shapiro 1964). The relative Shapiro delay between the images is

$$\Delta t_{\rm l} = \frac{4GM}{c^3}\ln\left(\frac{b_2}{b_1}\right), D_{\rm s}, D_{\rm ls} \gg b_{1,2}. \tag{14}$$

The impact parameters of the rays are $b_{1,2}$. This result is valid if b is much larger than the Schwarzschild radius of the lens, so the rays are only weakly deflected. The ray with smaller b is delayed more as it goes deeper into the lens' gravitational potential well. It also has a longer geometric path length (it forms the secondary image in Fig. 1).

Note this is the time delay at the lens. In reality, both photons must reach the Earth now, so the time of emission at the source must have been different. This requires an extra factor of the relative rates of a clock at the lens and at the source, $\frac{1+z_{\rm l}}{1+z_{\rm s}}$.

We then combined this Shapiro delay with the geometric path difference between the trajectories. Thus, the difference in time of emission required for photons traversing the two trajectories to reach the Earth simultaneously is

$$\Delta t_{\rm s} = \frac{1+z_{\rm l}}{1+z_{\rm s}}\frac{2GM}{c^3}\left[u\sqrt{u^2+4}+2\ln\left(\frac{u^2+u\sqrt{u^2+4}+2}{2}\right)\right]. \tag{15}$$

We used equation (10) to relate the source position u to the positions of its images. A correction for cosmological time dilation was also applied.

The change in redshift is given by the fractional difference in the scale factor of the Universe at the time of emission of the photons:

$$\Delta z = H(z_{\rm s})\Delta t_{\rm s}. \tag{16}$$

Using realistic parameters (Table 1 and $u \approx 1$), the DEE $\sim$1 m s^{-1}. In Section 3.1, we showed that the MCE is $\sim$1000 times larger, allowing us to neglect the DEE.

Even with more accurate instruments, a very large number of double image pairs would need to be observed before the random noise from source peculiar motions was reduced below such a small level. Thus, the DEE will not be important in the foreseeable future. An exception might possibly arise if the source galaxy peculiar motion could be estimated based on properties of galaxies near it.

4 DERIVATION OF THE DIFFERENTIAL MAGNIFICATION EFFECT FOR UNRESOLVED IMAGES

The effects mentioned in Section 3 cause the frequencies of identical photons emitted in different directions to end up different when measured at the Earth. The DME does not do this. It is merely an observational artefact due to inability to simultaneously resolve the images and take highly accurate spectra of them. This causes parts of the source with different redshifts to get blended together in the spectra. The precise way in which this blending occurs is different between the images.

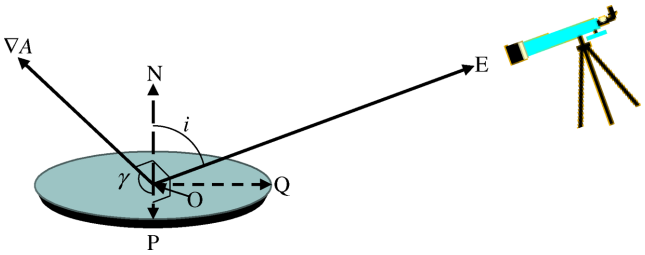


Figure 2. The observing geometry is shown here. The source galaxy has centre O and normal to its plane $\boldsymbol{ON}$. Earth is towards $\boldsymbol{OE}$, so the galaxy's inclination to the sky plane is i. $\boldsymbol{OQ}$ and $\boldsymbol{OP}$ are in the galaxy's plane and orthogonal to each other, with $\boldsymbol{OQ}$ as closely aligned with $\boldsymbol{OE}$ as possible. Thus, $\boldsymbol{OP}$ and $\boldsymbol{OE}$ are orthogonal. ∇A is directed within the source plane, so must also be orthogonal to $\boldsymbol{OE}$. ∇A is at an angle γ to $\boldsymbol{OP}$. The source is parametrized using cylindrical polar coordinates (r, ϕ), with centre O and initial direction $(\phi = 0)$ along $\boldsymbol{OQ}$.

We assume the spectra are integrated over the entirety of each image. The source is modelled as a typical spiral galaxy with exponential surface density profile and a realistic rotation curve. The lens is treated as a point mass. The parameters considered (Table 1) are designed with the Bullet Cluster (Tucker et al. 1995) in mind.

The basic idea is that spatially unresolved spectra can determine the intensity-weighted mean redshift of each image. This may be affected by rotation of the source galaxy. The effect is not reliant on an expanding Universe, so it will be simplest to think of the Universe as static for the remainder of this section.

The mean redshift velocity of each image is given by

$$\overline{v_{\rm r}} \equiv \frac{\int_{\rm Image} A\Sigma v_{\rm r}\,{\rm d}S}{\int_{\rm Image} A\Sigma\,{\rm d}S}. \tag{17}$$

The integrals are over area elements of the source S. This is treated as a disc with surface density

$$\Sigma = \Sigma_0\,{\rm e}^{-\frac{r}{r_{\rm d}}}. \tag{18}$$

The magnification A varies little over the source galaxy. This is because $\frac{r_{\rm d}}{D_{\rm s}} \ll \theta_{\rm E}$ (see Table 1). Thus, a linear approximation to A is sufficient:

$$A \approx A_0 + \frac{\partial A}{\partial u}{\rm d}u \quad (A_0 \equiv A \text{ at centre of source}). \tag{19}$$

In our model, A varies linearly with position in the source plane, but only in the direction directly away from the (projection of the) lens. In the orthogonal direction, A is independent of position at first order (because u is, and A depends only on u).

The geometry of the source is shown in Fig. 2. The radial velocity of any part of it is

$$v_{\rm r} = v_{\rm c}(r)\,\sin\phi\,\sin i. \tag{20}$$

Only the component of ∇A along $\boldsymbol{OP}$ is important. To see why, suppose that ∇A was entirely along $\boldsymbol{OQ}$. Reflecting the galaxy about the line OP without altering ∇A should reverse the DME as this is equivalent to reversing ∇A. However, the radial velocity of every part of the galaxy remains unaltered after the reflection (as $\phi \to \pi - \phi$). Thus, the DME must also remain unaltered.

Noting that the component of ∇A along $\boldsymbol{ON}$ is irrelevant for the DME, we see that only the component along $\boldsymbol{OP}$ might be relevant. This component causes the approaching and receding halves of the galaxy to be magnified differently. It will be responsible for the DME. Thus, we assume the lens is located somewhere along the line $\boldsymbol{OP}$, making ∇A entirely along this direction. The result is then multiplied by $\cos\gamma$.

The magnitude of the DME is therefore $\propto \cos\gamma \sin i$. Assuming isotropy, all values of γ are equally likely. But values of i close to $\frac{\pi}{2}$ are more likely because there are more ways for two vectors to be orthogonal than to be aligned. This means the ratio between the average magnitude of the DME and the maximum it could be is given by the mean of $|\cos\gamma \sin i|$, with γ unweighted but a further $\sin i$ weighting over i.[3] Thus,

$$\langle|\sin i\cos\gamma|\rangle = \frac{\int_0^{\frac{\pi}{2}}\cos\gamma\,\mathrm{d}\gamma}{\int_0^{\frac{\pi}{2}}\mathrm{d}\gamma}\frac{\int_0^{\pi}\sin^2 i\,\mathrm{d}i}{\int_0^{\pi}\sin i\,\mathrm{d}i} = \frac{1}{2}. \quad (21)$$

The angular separation between the lens and the unlensed source is given by

$$u = u_0 + \frac{\boldsymbol{r}.\widehat{\boldsymbol{OP}}}{D_\mathrm{s}\theta_\mathrm{E}} \quad (u_0 \equiv u \text{ at centre of source}). \quad (22)$$

The component of $\boldsymbol{r}$ (measured from the galaxy's centre) along $\boldsymbol{OP}$ is $r\sin\phi$. We have not kept careful track of signs because, for any conceivable orientation, the source galaxy could be rotating in the opposite sense, thereby reversing the DME. We explain which image has a lower redshift due to the DME later.

The difference in u between the centre of the source galaxy and any other point in it is given by

$$\mathrm{d}u = \frac{r\,\sin\phi}{D_\mathrm{s}\theta_\mathrm{E}}. \quad (23)$$

A_0 represents a constant magnification across the source. This does not contribute to the numerator in equation (17) because the radial velocity $v_\mathrm{r} \propto \sin\phi$, while Σ is independent of ϕ due to axisymmetry. Thus, integrating over ϕ gives 0.[4] The DME arises when including the first-order correction to A.

The denominator in equation (17) is a normalization for each image (its total intensity).[5] Because the magnification is nearly constant across the source galaxy, we can approximate that $A = A_0$. The first-order correction to A would have a $\sin\phi$ dependence, which is irrelevant when integrated over all ϕ. This further justifies our approximation. Therefore, the denominator in equation (17) becomes

$$\int_\mathrm{Image} A\Sigma\,\mathrm{d}S = \Sigma_0\pi{r_\mathrm{d}}^2\left(\frac{u^2+2}{u\sqrt{u^2+4}} \pm 1\right). \quad (24)$$

To understand the sign of the DME, first note that regions closer to the lens are magnified more. In our approximation, the numerator in equation (17) is determined by $\frac{\partial A}{\partial u}$, which is the same for both images (equation 12). Therefore, the image with the lower magnification (the secondary image) has a larger $|\overline{v_\mathrm{r}}|$. Thus, if it was known which side of the rotating source was the approaching side, one could determine which image should have a higher mean redshift due solely to the DME.

Including the second-order dependence of A on sky position slightly alters the calculations done so far. Because a second-order term does not affect the approaching and receding halves of the source galaxy differently, the numerator in equation (17) is unaltered. But the denominator is affected, because the total intensity of each image may be altered by a second-order term. This means that our derivation of the DME has a fractional error which is second order in $\frac{r_\mathrm{d}}{D_\mathrm{s}\theta_\mathrm{E}}$. We consider this acceptable and proceed to develop a model for the redshift structure of the source. This requires a rotation curve.

4.1 Model rotation curves

It will likely be difficult to directly observe the source galaxy rotation curve $v_\mathrm{c}(r)$ as it is very far away. It is also difficult to precisely determine its surface density and thus predict the form of $v_\mathrm{c}(r)$. Fortunately, we are considering a disc-integrated spectrum and so the exact shape of $v_\mathrm{c}(r)$ will turn out not to be very important once the maximum level v_max is fixed.

To get a rough idea of $v_\mathrm{c}(r)$, we take advantage of the tight empirical relation between the forces in rotating disc galaxies as required to sustain their rotation curves and those predicted by Newtonian gravity based on the visible (baryonic) mass (Famaey & McGaugh 2012, and references therein). This empirical formalism goes by the name of modified Newtonian dynamics (MOND; Milgrom 1983). Regardless of whether it is correct at a fundamental level, it does seem to provide a good empirical way of predicting rotation curves. Here, this is important because measuring the actual rotation curve of the source galaxy would be very challenging.

The particular empirical relation we adopt follows the work of Famaey & Binney (2005):[6]

$$\left(\frac{|\boldsymbol{g}|}{|\boldsymbol{g}|+a_0}\right)\boldsymbol{g} = \boldsymbol{g}_\mathrm{N}, \quad (25)$$

where $\boldsymbol{g}$ is the true gravitational field strength while $\boldsymbol{g}_\mathrm{N}$ is the prediction of Newtonian gravity based on the visible mass. a_0 is an acceleration scale ($\approx 1.2\times10^{-10}\,\mathrm{m\,s^{-2}}$) below which either gravity becomes non-Newtonian or DM must be considered in addition to the baryons. Thus, the magnitude of the gravitational field is given by

$$g = \frac{g_\mathrm{N}}{2} + \sqrt{\left(\frac{g_\mathrm{N}}{2}\right)^2 + g_\mathrm{N}a_0}. \quad (26)$$

It is not worthwhile to accurately determine g_N for any particular mass distribution because the actual mass distribution in the source is uncertain. Thus, we approximated g_N using an analytic method. We assumed that, to determine g_N at a particular in-plane location, only material at smaller radii need be considered (we verified that the force from material at larger radii was very small). The Newtonian force at a distance r from the centre of a narrow ring of mass $\mathrm{d}M$ and radius x is

$$g_\mathrm{N} \approx \frac{G\,\mathrm{d}M}{r^2} + \frac{3G\,\mathrm{d}M\,x^2}{4r^4} \quad (x < r, \text{ interior ring}). \quad (27)$$

[3] Edge-on galaxies are less likely to be detected due to dust obscuration. This makes low values of i – and thus a smaller DME – more likely, for a randomly selected multiple image.

[4] This is expected, as the DME does not arise if the image is uniformly magnified.

[5] What we perceive as the total intensity given D_s and z_s, but without correcting for magnification by the lens.

[6] This is the so-called 'simple μ-function' in MOND.

This is correct at second order in $\frac{x}{r}$. The total force at any point within the disc was found by decomposing the galaxy into a large number of rings with $\mathrm{d}M = 2\pi x\,\mathrm{d}x\,\Sigma(x)$. We then summed only the forces resulting from interior rings. Therefore,

$$g_{\mathrm{N}} = \int_0^r \left(\frac{G\,\mathrm{d}M}{r^2} + \frac{3G\,\mathrm{d}M\,x^2}{4r^4}\right) \tag{28}$$

$$= 2\pi G\Sigma_0 f(\widetilde{r}), \quad \text{where} \tag{29}$$

$$\widetilde{r} \equiv \frac{r}{r_{\mathrm{d}}} \quad \text{and} \tag{30}$$

$$f(\widetilde{r}) = \frac{1-\frac{13}{4}\,\mathrm{e}^{-\widetilde{r}}}{\widetilde{r}^2} - \frac{7\,\mathrm{e}^{-\widetilde{r}}}{4\widetilde{r}} + \frac{9\left(1-\mathrm{e}^{-\widetilde{r}}-\widetilde{r}\,\mathrm{e}^{-\widetilde{r}}\right)}{2\widetilde{r}^4}. \tag{31}$$

When obtaining the true value of g from g_{N}, the ratio $\frac{g_{\mathrm{N}}}{a_0}$ is important. Therefore, we introduce a new variable, the dimensionless density k,

$$k \equiv \frac{G\Sigma_0}{a_0}. \tag{32}$$

Typical values for k are order 1. Using the empirical equation (26) to get g from g_{N}:

$$\frac{g}{a_0} = \pi k f(\widetilde{r}) + \sqrt{\left(\pi k f(\widetilde{r})+1\right)^2 - 1}. \tag{33}$$

To get the rotation curve, we equate g with the centripetal acceleration. Thus,

$$\frac{v_{\mathrm{c}}^2}{\widetilde{r}\,r_{\mathrm{d}}} = \left(\frac{g}{a_0}\right) a_0, \tag{34}$$

$$v_{\mathrm{f}} = \sqrt[4]{2\pi k}\sqrt{r_{\mathrm{d}} a_0}. \tag{35}$$

The rotation curve flat lines at the level $v_{\mathrm{f}} = \sqrt[4]{GMa_0}$, where the total disc mass $M = 2\pi\Sigma_0 r_{\mathrm{d}}^2$. The shape of the rotation curve is given by

$$\widetilde{v_{\mathrm{c}}}(\widetilde{r}) \equiv \frac{v_{\mathrm{c}}(\widetilde{r})}{v_{\mathrm{f}}} = \frac{\sqrt{\pi k\widetilde{r} f(\widetilde{r}) + \widetilde{r}\sqrt{\left(\pi k f(\widetilde{r})+1\right)^2 - 1}}}{\sqrt[4]{2\pi k}}. \tag{36}$$

We show two such rotation curves in Fig. 3. Here, we also show how K affects the ratio between v_{max} and v_f.

4.2 The final result

Combining our results, we get that

$$|\overline{v_{\mathrm{r}}}| = \frac{v_{\mathrm{f}} r_{\mathrm{d}} \sin i \cos\gamma}{D_{\mathrm{s}}\theta_{\mathrm{E}}} \times \frac{\int_0^\infty \int_0^{2\pi} \overbrace{\mathrm{e}^{-\widetilde{r}}}^{\propto\Sigma}\, \widetilde{v_{\mathrm{c}}}(\widetilde{r})\widetilde{r}^2\, \overbrace{\frac{4}{u^2(u^2+4)^{\frac{3}{2}}}}^{-\frac{\partial A}{\partial u}} \sin^2\phi\,\mathrm{d}\phi\,\mathrm{d}\widetilde{r}}{\underbrace{\pi\left(\frac{u^2+2}{u\sqrt{u^2+4}} \pm 1\right)}_{\propto A}}. \tag{37}$$

The magnification A changes by order 1 over an angular distance of θ_{E}, while the angular radius of the source galaxy $\sim\frac{r_{\mathrm{d}}}{D_{\mathrm{s}}}$. Thus, the DME as a fraction of the typical radial velocity of the source is $\sim\frac{r_{\mathrm{d}}}{D_{\mathrm{s}}\theta_{\mathrm{E}}}$. The galaxy's typical radial velocity is $v_{\mathrm{f}}\sin i$. Another factor of $\cos(\gamma)$ is needed to account for the lensing geometry. As can be seen from equation (37), this provides a rough guide to the DME (as $u \sim 1$ for a realistic target).

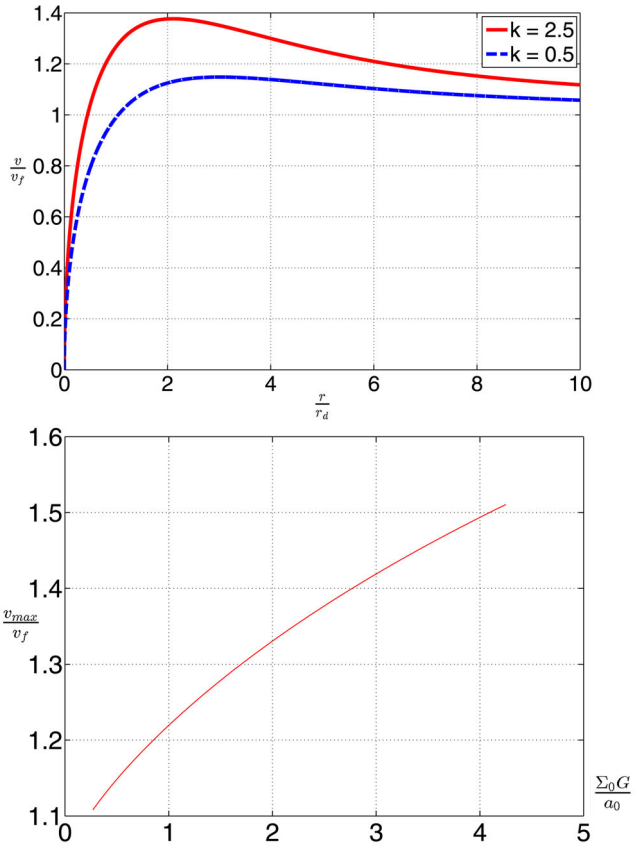


Figure 3. Top: rotation curves resulting from equations (31) and (36), used in this work. $v_{\mathrm{c}}(r)$ flat lines at v_{f}. The surface density $\Sigma = \Sigma_0\,\mathrm{e}^{-\frac{r}{r_{\mathrm{d}}}}$. The parameter k controls the shape of the rotation curve ($k \equiv \frac{\Sigma_0 G}{a_0}$). Bottom: the ratio of maximum to flat line rotation speed as a function of central surface density.

An important quantity for the DME is the difference in $\frac{1}{A}\frac{\partial A}{\partial u}$ between the images:

$$\Delta\left(\frac{1}{A}\frac{\partial A}{\partial u}\right) = \frac{\partial A}{\partial u}\Delta\left(\frac{1}{A}\right) = \frac{4}{\sqrt{u^2+4}}. \tag{38}$$

In equation (37), the integration over ϕ yields π. The integral over r is not analytic. Thus, we define

$$I \equiv \int_0^\infty \mathrm{e}^{-\widetilde{r}}\,\widetilde{v_{\mathrm{c}}}(\widetilde{r})\,\widetilde{r}^2\,\mathrm{d}\widetilde{r}. \tag{39}$$

Substituting for θ_{E} using equation (8), we get that

$$\Delta\overline{v_{\mathrm{r}}}|_{\mathrm{DME}} = \frac{v_{\mathrm{f}}\, r_{\mathrm{d}}\, \sin i\cos\gamma\, I\, c\sqrt{D_{\mathrm{l}}}}{\sqrt{u^2+4}\,\sqrt{GMD_{\mathrm{ls}}D_{\mathrm{s}}}}. \tag{40}$$

The integral I depends somewhat on the central surface density in the sense that, for the same v_{f}, the DME is greater at higher k. However, the *maximum* rotation speed is very well correlated with the DME. In fact, the ratio $\frac{I}{\widetilde{v}_{\max}} = 1.89 \pm 0.02$ for $k = 0.1 \to 5$. As maximum rotation speeds are generally easier to determine than the flat line level, this makes correcting for the DME easier.

If the surface density declines sufficiently slowly with r, then the integral I might diverge. This is due to limited validity of a linear

approximation to A – a more careful treatment would be required. This might apply to rotating elliptical galaxies with $\rho \propto r^{-4}$. But even if $\Sigma \propto r^{-3}$, the divergence of I is fairly slow. Thus, although the integral would need a cut-off radius, its precise value would not affect the result much.

A linear approximation to A must break down if u changes by order 1. Thus, a logical cut-off might be the Einstein radius (at the source plane) or the distance between the source and the projected lens.

If a fibre-fed spectrograph was used or the field of view was otherwise restricted, then this may impose an obvious cut-off. For an inclined disc galaxy, a circular aperture would cover a non-circular region in the disc plane. One might need to take this into account.

We now decide on realistic parameters to gain a feel for the scale of the DME and MCE. Ideally, one would like to measure the motion of both components of the Bullet Cluster. However, we choose a mass corresponding roughly with the subcluster in the Bullet (Mastropietro & Burkert 2008). This is because the centre of mass likely has little peculiar velocity as there is little structure on such large scales. Thus, the lower mass component probably moves faster (with respect to the Hubble flow).

The MCE $\propto v\sqrt{M}$ (equation 13). Assuming also that $v \propto \frac{1}{M}$ for the components of the Bullet Cluster and that the value of u would be broadly similar whichever component is targeted, the MCE overall $\propto \frac{1}{\sqrt{M}}$. This makes it larger around the subcluster. Furthermore, using its motion to extrapolate the total collision velocity is much more reliable than using the motion of the main cluster, because the subcluster contributes most of the relative velocity.

A typical source galaxy orientation is chosen using equation (21). Although we used $v_\mathrm{f} = 100$ km s^{-1}, it is around double this for our own Galaxy (e.g. McMillan 2011).

Using the parameter values in Table 1, we obtained the results shown in Fig. 4. The DME and MCE are comparable in magnitude.

Observing a similar source multiply imaged by the higher mass component instead does not reduce the relative importance of the DME. This is because the DME $\propto \frac{1}{\sqrt{M}}$ (equation 40), just like the MCE. In fact, this scaling highlights an additional problem: substructure in the lens (e.g. individual galaxies) with much less mass than the entire cluster can enhance the DME. For example, an elliptical galaxy in the lens plane with $M = 10^{13}\,\mathrm{M}_\odot$ would cause a DME ∼3 times larger than the smooth cluster potential.

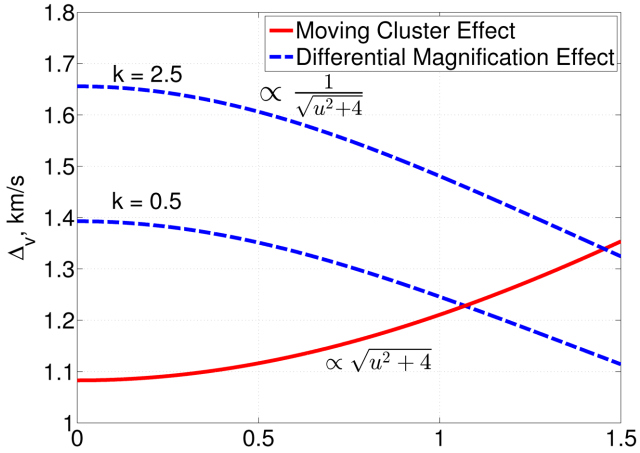


Figure 4. The difference in redshift between double images of a typical background galaxy as a function of its position, due to the effects described in the text (equations 13 and 40). Parameter values used are listed in Table 1. The shape of the rotation curve (k) has a modest impact on the DME once its flat line level v_f is fixed. If instead v_max is held fixed, then the impact of k on the DME is very small.

This problem could be mitigated to some extent by not selecting images which show indications of being lensed by small-scale structure (e.g. avoiding images appearing near a galaxy in the lens plane). We have implicitly assumed that such a selection has been done, such that a point mass model for the lens is appropriate. Even in this case, it might well be necessary to correct for the DME. This correction would be less relevant if targets could be selected for which the effect is small. We now consider how these things might be achieved.

5 CORRECTING FOR THE EFFECT

For spatially unresolved spectra, it is possible to calculate the DME by determining the parameters in equation (40). If radial velocities accurate to a few km s^{-1} are obtained for a galaxy, then determining $v_\mathrm{max} \sin i$ to ∼10 km s^{-1} should be feasible using widths of spectral lines (see later).

r_d might be obtained from an image of the target, once distortion and magnification by the lens were corrected for. If the image were taken at more than one wavelength, it would suggest a value for k (which we do not need very accurately) as the colour can be used to estimate the baryonic M/L.

There is no need to determine the inclination as we are only interested in redshift gradients across the source. However, the orientation of the major axis of the image is important in determining the axis of rotation and thus the angle γ in Fig. 2.

To know the sense of rotation (i.e. which side of the source galaxy is the approaching side), we would need spectra of different parts of the source galaxy. Naturally, a disc-integrated spectrum would be insufficient for this purpose. However, one could make do with poorer spectral resolution.

The secondary image is inverted relative to the primary, providing an important consistency check if both images were used for such a determination. We strongly recommend doing this, because an error would lead to a 200 per cent error in the calculation of the DME. The chance of this is minimized with two determinations of the sense of rotation.

Finally, we also need ∇A, which must come from a lensing reconstruction.

5.1 Additional information from detailed spectral line profiles

It is often possible to extract more information from a spectral line than just the location of its centroid. The width of the line profile can be used to estimate e.g. $v_\mathrm{max} \sin i$.

The MCE simply shifts the entire spectrum. The DME leads to a 'tilt' being introduced because one side of the galaxy is magnified more than the other. These effects are different. Therefore, detailed line profiles can tell us if the shift in the centroid of spectral lines is due to the MCE or the DME. This would avoid the need to determine parameters like the disc scale length and orientation. A detailed lensing reconstruction to determine ∇A would also be avoided.

We investigate how the DME and MCE affect line profiles of disc galaxies with rotation curves parametrized by equation (36). We assume the galaxy is viewed edge-on, so the radial velocity of any part of it is

$$v_\mathrm{r}(r, \phi) = v_\mathrm{c}(r) \sin(\phi). \quad (41)$$

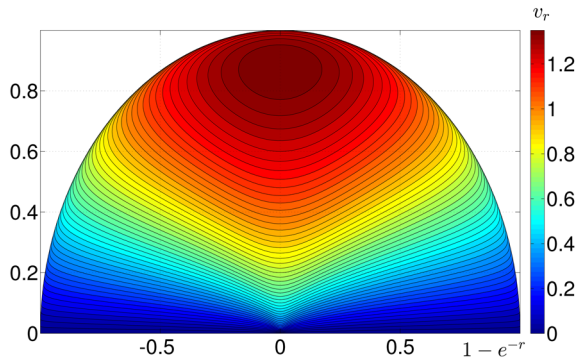


Figure 5. Radial velocity map of a disc galaxy viewed by an observer within its plane at large x, for the case $k = 2.5$. Radial velocities are antisymmetric about the x-axis. The radial coordinate is rescaled so all parts of the figure would be equally bright. The units are such that $r_d = 1$ and $v_f = 1$. Note the large region with v_r close to the maximum value. The result for $k = 0.5$ is very similar, although v_{max} is much closer to v_f.

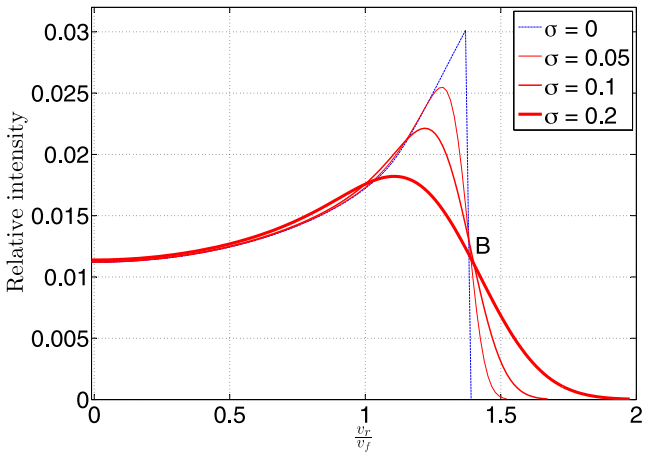


Figure 6. The synthetic line profile of an intrinsically narrow line in an unlensed galaxy with $k = 2.5$, viewed edge-on. The profile is symmetric about $v_r = 0$. Velocities are scaled to the flat line level v_f. The sharp drop in the line profile (blue) would probably get blurred (e.g. by random motions), so we convolved the profile with Gaussians of widths σ. The results are shown as red lines with thicknesses $\propto \sigma$. Notice how all four profiles pass close to the point marked B. The result for $k = 0.5$ is similar, if the profiles are scaled to have the same v_{max} rather than v_f.

The resulting radial velocity map is shown in Fig. 5. Only half of the galaxy is shown because v_r is antisymmetric about the viewing direction (the x-axis). v_r is symmetric about the y-axis, because $\sin\phi = \sin(\pi - \phi)$.

To determine the profile of a narrow spectral line, we divide the galaxy up into a large number of elements. We use cylindrical polars so v_r becomes separable. Thus, the rotation speed only needs to be calculated once at each r (for all ϕ). The radial velocity at the centre of each element is used to classify it among 200 bins in radial velocity.

Assuming constant mass-to-light ratio (M/L) for the baryons, the total intensity of the element multiplied by the magnification A is then assigned to the corresponding radial velocity bin. Because radial velocities and wavelengths are directly related, in this way one obtains a synthetic line profile.

Spectral lines have an intrinsic width and can be further broadened by random motions within the galaxy. To account for these effects and also for instrumental errors, we convolved our synthetic line profiles with Gaussians of various widths σ. The results are shown in Fig. 6. The sharp peaks at $v_r \approx \pm v_{max} \sin i$ give rise to the name of a double-horned profile.

These horns are caused by the rotation curve having a peak, leading to a small range in v_r corresponding to a large range in r. The greatest attained values of $|v_r|$ also correspond to large ranges in ϕ, because $\sin\phi$ is nearly independent of ϕ when $\phi \approx \pm\frac{\pi}{2}$.

Thus, a small range in v_r corresponds to a large region in the galaxy. Moreover, the peak rotation speed occurs at a radius close to that which maximizes the light emitted per unit radius ($r = r_d$). Fig. 5 shows the 'bull's eye' corresponding to the fairly large region with near-maximal $|v_r|$. This is responsible for the very pronounced horns in the line profile. They are somewhat less pronounced at high σ.

Although one might expect a feature corresponding to v_f (at least at low σ), this is absent. A quick look at Fig. 5 shows why $v_c(r) \approx v_f$ only for sufficiently large r. There is very little light from such regions, so a disc-integrated spectrum is hardly sensitive to them. In fact, due to the steep decline in surface brightness with r, most of the spectral intensity at $v_r = v_f$ actually comes from the rising part of the rotation curve (when $v_c \sin\phi = v_f$) rather than from the flat part. Thus, in the line profile, there is nothing special about v_f. This is not true for v_{max}.

Determining $v_{max} \sin i$ from a line profile is non-trivial as the horns move to lower $|v_r|$ as σ increases. Instead of using the horn positions, one could use the values of v_r where the intensity is a certain fraction of the intensity at the line centre ($v_r = 0$). If this fraction is chosen carefully, then one could simply scale the resulting v_r by a constant factor and accurately recover $v_{max} \sin i$ over a wide range in σ and k. To see why, note that spectra with different σ all pass close to the point marked B in Fig. 6.

Ultimately, it might be better to compare the observed line profile with a suite of synthetic profiles built for a range of k, σ and $v_{max} \sin i$. The initial guess for σ might come from considering the shape of the tail. If $v_{max} \sin i$ is accurately recovered, then the DME hardly depends on k.[7]

The horns are cause by a relatively small part of the galaxy but they greatly affect the mean radial velocity of its image. Thus, if the galaxy was not axisymmetric and e.g. had a dusty spiral arm obscuring light from this region, then the redshift measurement of each image may be biased. Partly for this reason, it may be a good idea to consider the rest of the line profile and not just the mean redshift (which is basically the same as considering just the horns).

We now consider how the DME and MCE affect the line profile. The mean redshift velocity of the line is raised by 1 per cent of v_{max} (1.08 per cent of v_f for $k = 0.5$ and 1.26 per cent for $k = 2.5$). We consider separately the cases where either the DME or the MCE is wholly responsible for this shift in line centroid. We also construct control line profiles like those in Fig. 6. These are obtained by setting

$$A = 1 \; \forall_{r,\phi}. \tag{42}$$

[7] Line profiles can also be used to find $v_f \sin i$ without detailed rotation curves. In this case, the value of k is important as a 'Σ correction' must be applied to get from v_{max} to v_f (e.g. bottom panel in Fig. 3). k does not affect the line profile much and so it would need to come from an image and photometry.

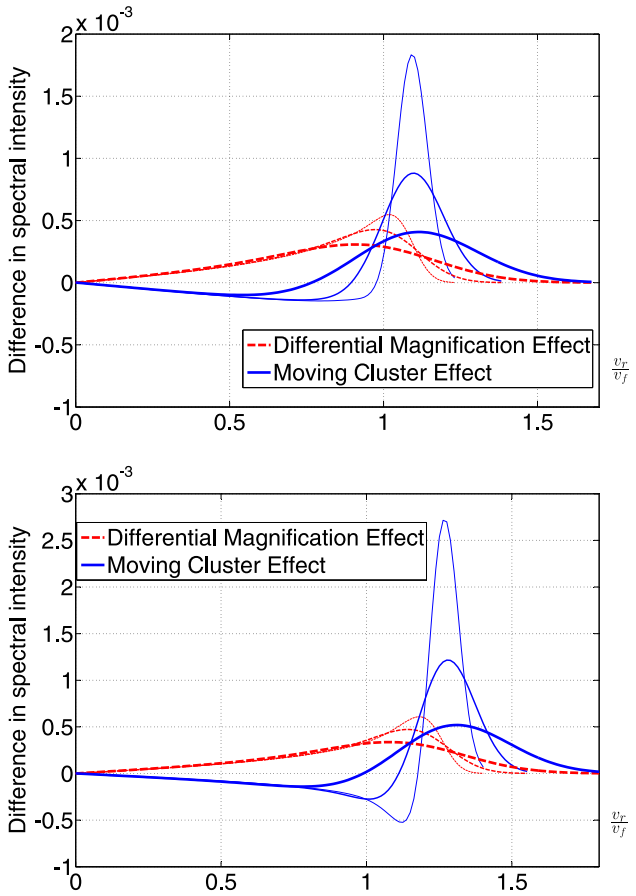


Figure 7. The residuals in the spectral profile due to the DME (equation 43) and the MCE (horizontal shift of profile) are shown here. These were obtained by subtracting a control line profile (equation 42). The patterns are antisymmetric about $v_r = 0$. Results are for an edge-on galaxy with $k = 0.5$ (top) and $k = 2.5$ (bottom). Both effects change the mean redshift by 1 per cent of the maximum rotation speed, representing 1.08 per cent of v_f for $k = 0.5$ and 1.26 per cent for $k = 2.5$. The spectra were convolved with Gaussians of widths 0.05, 0.1 and 0.2 v_f (higher σ indicated by thicker line). The MCE cannot change the amplitudes of the horns. The DME makes one more pronounced and the other less.

In Fig. 7, we show the pattern of residuals (relative to the control) created by each effect. The total line intensity is kept the same for the comparison.

To obtain the corresponding observations, one would need to account for the images having different overall magnifications. Thus, the spectra would have to be rescaled. We assume this can be done perfectly (i.e. the photometry is very accurate).

The MCE corresponds to a horizontal shift in the spectrum relative to the control. This means the amplitudes of the horns are unaffected. *The pattern of residuals corresponds to the gradient in the spectrum.* Thus, the residuals are largest near the positions of the horns, but small precisely at them. The residuals are of opposite signs on either side of each horn because the gradient in the spectrum changes sign there.

For the DME, we set

$$A = 1 + n\widetilde{r}\sin\phi. \tag{43}$$

Note that no assumptions are made about any of the factors controlling the amplitude of the DME, beyond it being a small effect relative to $v_{max}\sin i$ (i.e. $n \ll 1$) and that we need not consider the second-order dependence of A on position in the source plane. The purpose here is to illustrate how the DME affects the line profile, not how much (this is controlled by n). If the DME $\sim 0.01 v_f \sin i$, then the residuals would be $\sim$1 per cent of the line profile.

The image overall is not magnified for any (small) n. We adjust n until the line centroid shifts by the correct amount, to allow comparison with the MCE.

The DME causes one side of the galaxy to be magnified more than the other. Thus, the residuals due to it are of the same sign for each half of the galaxy (e.g. for $v_r > 0$). There is no change in sign at the horns.

These correspond to material displaced from the centre of the galaxy along the direction $\boldsymbol{OP}$ in Fig. 2. As argued previously, we only need to consider the component of ∇A along this direction. Thus, the effects of differential magnification are substantial for the material corresponding to the horns (in so far as the DME affects the image at all). This is in contrast to the MCE, which hardly affects the line profile at these positions (because the gradient of the line profile there is 0).

For some v_r, the MCE leads to very large residuals if $\frac{\sigma}{v_f \sin i}$ is low (Fig. 7). Thus, observing faster rotating galaxies might make it easier to distinguish between the DME and the MCE (as $\frac{\sigma}{v_f \sin i}$ would likely be smaller). However, the DME would be larger and so it would have to be accounted for more accurately.

Detailed profiles of individual spectral lines may therefore help in determining the balance between the MCE and the DME in accounting for redshift differences between multiple images. In reality, a large number of spectral lines would probably need to be stacked. Even then, it seems likely that, in so far as redshift differences between the images are discernible, the cause of such differences can also be determined.

5.2 The second-order effect

For simplicity, we continue assuming the source is located along $\boldsymbol{OP}$. Thus, regions with high $|v_r|$ are magnified more and regions with lower $|v_r|$ are magnified less than the centre of the source due to the second-order dependence of A on position. To investigate what this means for spectral line profiles, we set A to depend quadratically on position along the direction $\boldsymbol{OP}$. This introduces a $\sin^2\phi$ dependence:

$$A = \frac{1 + n\widetilde{r}^2 \sin^2\phi}{1 + 3n}. \tag{44}$$

A quadratic dependence along the orthogonal direction would give a $\cos^2\phi$ dependence. Because $\cos^2\phi = 1 - \sin^2\phi$, a second derivative of A in either direction would affect the line profile in the same way (i.e. the residuals would have the same pattern, up to sign); once any overall magnification was corrected for.

When comparing the spectra, observers would first scale them to have equal intensities. Thus, we must avoid changing the total intensity. This leads to the factor of $1 + 3n$ in the denominator of equation (44).

The results are shown in Fig. 8. The effect is symmetric in v_r, so both horns are equally affected. These end up more pronounced in the secondary image than in the primary (in this example).

In reality, both the first- and second-order DME would be present for any given pair of images of the same object. Thus, the residuals would have both an antisymmetric and a symmetric part. However, the latter would likely be very small for cluster mass lenses (except for caustic images).

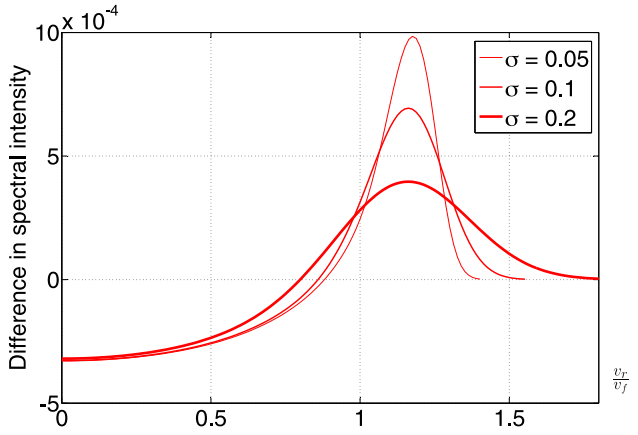


Figure 8. The pattern of residuals for the second-order DME (equation 44) and $k = 1$. A control profile obtained using equation (42) was subtracted and the result convolved with a Gaussian. The residuals are symmetric about $v_r = 0$, so both horns become more pronounced in this example.

5.2.1 Non-rotating sources

We briefly consider how the DME might affect a non-rotating pressure-supported galaxy, such as an elliptical. If a galaxy is symmetric such that $\rho(\boldsymbol{r}) = \rho(-\boldsymbol{r})$ and $\sigma(\boldsymbol{r}) = \sigma(-\boldsymbol{r})$, then at first order the DME does not affect an unresolved image at all. To see this, consider an inversion mapping $\boldsymbol{r} \rightarrow -\boldsymbol{r}$ while leaving ∇A unchanged. The situation is identical to reversing the direction of ∇A instead, so one expects the DME to act in exactly the opposite way on the spectrum. But the situation has not physically changed, so the DME must also remain unchanged.

This conclusion breaks down at second order. Suppose parts of the galaxy further from its centre are magnified more. Then, as the velocity dispersion generally decreases outwards, the derived velocity dispersion of the image will be reduced by the DME. The effect is larger for the fainter (secondary) image, which will thus appear to have a smaller velocity dispersion than the primary in this example.

This is likely to be more important for galaxy–galaxy lensing as $\theta_{\rm E}$ is smaller, making du over the source larger. In this case, the DME might be useful to constrain the form of $\sigma(r)$ using a double image of a distant elliptical galaxy.

Alternatively, if the source galaxy was well understood, one might be able to constrain ∇A and thus have a better understanding of the lens. Doing both simultaneously would likely be very challenging and model dependent.

6 TARGETS WITH A SMALLER EFFECT

Fig. 4 shows that the DME may well need to be accounted for when using the redshifts of double images to determine the motion of the lens. However, doing this accurately may be difficult because of the cosmological distance to the source galaxy. Therefore, we suggest sources for which the DME should be smaller, allowing us to correct for it less accurately.

Some strategies outlined here involve selecting targets which are harder to observe, thereby making their spectra less accurate. It is up to observers to decide which targets best minimize the uncertainty introduced by the DME while still being feasible to obtain accurate spectra for. We also note that minimizing the uncertainty introduced by correcting for the DME is not necessarily equivalent to minimizing the magnitude of the DME, because there may be sources for which the DME can be estimated more reliably.

6.1 QSOs

The DME $\propto rv$, where the source has typical size r and radial velocity spread v. For a given mass M, the Virial theorem gives $v \propto \frac{1}{\sqrt{r}}$. Thus, the DME $\propto \sqrt{r}$. For sources with a particular M, the DME would be reduced if the source were smaller, even though it would spin faster.

One obvious type of very small target visible over cosmological distances is a quasi-stellar object (QSO). If a doubly imaged QSO could be found lensed by the Bullet Cluster, it might make an excellent target.

QSO spectra can sometimes lack distinctive features which are required for precise redshift measurements. The Lyα forest might provide a solution, but only if the same feature appeared in both images. Because the rays of light corresponding to the images diverge from the source,[8] this is only feasible if the gas cloud causing the absorption feature was located fairly close to the QSO.

Another problem might be that the small size of QSOs makes their radiation time variable. Thus, the time delay between the images could make it difficult to compare their spectra. This might require observers to wait out the time delay, which would first have to be determined (though it could be estimated, perhaps using equation 15).

6.2 Smaller and fainter galaxies

The DME is proportional to both the rotation velocity and the size of the source galaxy. Brighter galaxies generally rotate faster (Tully & Fisher 1977), so targeting fainter galaxies might help. One advantage of this approach is that the number density of fainter galaxies is greater than for brighter ones (Schechter 1976). This makes it more likely that suitably oriented multiple images can be found.

However, it would be harder to obtain accurate spectra – and thus redshifts – for fainter targets. Given the high accuracy required in the redshift measurements and the cosmological distance to the source, this is perhaps not the best option at present.

6.3 Elliptical galaxies

Elliptical galaxies might make good targets as they usually rotate slower than spirals, if at all. They might be distinguished using colour or image shape (though one might need to correct for distortion by the lens). The surface brightness declines outwards much more gradually for ellipticals than for spirals, potentially providing another way of finding them.

Before conducting detailed observations, targets selected like this might be followed up to check if the spectral line profiles were double horned (characteristic of rotation along the line of sight). A good target should have a Gaussian-looking line profile, characteristic of a pressure-supported object.

However, even ellipticals can rotate, so the DME might not be eliminated by observing one. Also, most galaxies are not elliptical, so finding a bright doubly imaged one is somewhat dependent on luck. Nonetheless, we consider this the best option. This is partly because the work of Gonzalez et al. (2009) identified a multiply imaged galaxy which may be a good target for determining the MCE.

[8] By an angle $\frac{1+z_{\rm l}}{1+z_{\rm s}} \frac{D_{\rm l}}{D_{\rm ls}} (\theta_1 - \theta_2)$.

6.4 Galaxy orientation and viewing angle

Fig. 2 shows the geometry of the situation. The radial velocity of any part of the galaxy is scaled by $\sin i$, so a face-on spiral could not have a redshift gradient across it and thus would be unaffected by the DME.

Determining i requires an image of the source galaxy to determine its shape. One could imagine trying to select targets which look round. Even then, the target might be an elliptical galaxy with some rotation along the line of sight.

The direction of ∇A is also very important. In theory, we should seek situations where ∇A is orthogonal to the major axis of the image. In such situations, an edge-on disc galaxy would appear as a line on the sky aligned orthogonally to the direction towards the lens. With more complicated lenses, the galaxy would appear as a line on the sky orthogonal to ∇A, which hopefully could be estimated using a lensing reconstruction.

6.5 Galaxy position and caustics

As was already pointed out in Molnar et al. (2013a), the MCE is maximal for image separations aligned with the direction of the lens' proper motion. As the collision is nearly in the plane of the sky, the likely direction of this motion is plain to see and observers should target double images separated approximately along this direction.

In the simple lens model that we use, the DME $\propto \frac{1}{\sqrt{u^2+4}}$ while the MCE $\propto \sqrt{u^2+4}$. Thus, galaxies less closely aligned with the lens make better targets in terms of the systematic error of the DME. For such sources, the images are more widely separated.

Unfortunately, sources with larger u make worse targets under a number of other considerations. Both images – but especially the secondary – are fainter. This image also becomes very close to the lens, making it more likely to be obscured.

A lensing reconstruction could be used to suggest particular locations where the magnification is nearly constant. A galaxy with double images near such locations might make a good target for measuring the MCE. The difficulty with this is that such 'sweet spots' might be small and not have any observable galaxies in them. Also, a magnification map of sufficient accuracy might be difficult to obtain.

Regions where A varies rapidly with position enhance the DME. Caustics occur where the magnification of a small part of the source plane is infinity. This means that the magnification varies rapidly with position in the source plane, greatly increasing the DME. For this reason, it has been suggested to avoid caustic images (Molnar et al. 2013a).

However, it may be worthwhile to try correcting for the DME in caustic images because they are generally very bright, making for more accurate spectra. The correction would need to be done very accurately in this case, because the MCE might be much smaller than the DME. Because this is likely to lead to controversy surrounding the measurements, we also recommend avoiding caustic images unless the observational case is compelling.

6.6 Substructure within the source galaxy

If the source galaxy has e.g. a bright star-forming region which emits strongly in the ultraviolet (UV) while the rest of the galaxy does not, then another possibility arises. UV spectral lines would correspond to material in a small part of the galaxy. Consequently, different parts of it would have much the same radial velocity and the magnification across it would be more uniform than across the whole galaxy. This would reduce the DME for the UV lines.

Thus, in this example, the redshifts for the images should be calculated using only the UV lines. In practice, one would exploit the fact that a small part of the galaxy should have only a narrow range of redshifts. Thus, one might use only spectral lines which have a similar redshift. If the intrinsic linewidth was small, then the line should be very narrow as there would not be much rotational broadening.

Another possibility is using spectral lines that are more prominent in the bulge of the galaxy (if it has one). The bulge is mostly pressure supported with little rotation and is also much smaller than the whole galaxy. In this case, the spectral lines to use might be quite broad, but have a Gaussian line profile even if the galaxy is rotating (so most spectral lines have a double-horned profile).

For this technique to be of much benefit, the galaxy needs to be quite inhomogeneous in some way. It might be difficult to tell whether it is from an image. Also, the technique reduces the number of spectral lines that are used to calculate the redshift, making for less precise measurements. This makes it difficult to target fainter galaxies, perhaps forcing observers to choose between observing all of a fainter galaxy or (effectively) part of a brighter one.

Any decision to restrict which spectral lines are used to determine the MCE should be justified based on more detailed observations of nearby galaxies. This increases confidence that the decision does indeed effectively restrict the observations to a small part of the source.

Even if all usable spectral lines are used to measure the MCE, it is still likely that some lines are less affected by the DME than others. It may be important to allow for this in the analysis e.g. by grouping spectral lines based on their linewidth and shape and obtaining an inference on the MCE for each group.

Because of the uncertainties introduced by such procedures, we recommend reducing the DME by careful choice of target so that the exact method used to correct for it does not much affect the inferred lens velocity.

7 OBSERVATIONAL PROSPECTS

We now consider the technical feasibility of detecting the MCE with high-resolution spectroscopic measurements. The target we consider is presented in Gonzalez et al. (2009). This has a flux of $\sim$100 mJy at mm wavelengths. Because of dust in the source galaxy, it is best to do the observations at such wavelengths. For this purpose, we consider using ALMA. The Bullet Cluster rises to within $\sim 35^\circ$ of the zenith at this site (minimum airmass $\approx$1.2).

The parameters we supplied are given in Table 2. To clarify the tension with ΛCDM, it would be necessary to constrain the collision

Table 2. Input parameters used for the ALMA exposure time calculator, available at https://almascience.eso.org/proposing/sensitivity-calculator. The dual polarization mode should be used as polarization is unimportant here. The angular resolution does not affect the result, which was 6.17 h.

Parameter	Value
Declination	-56°
Frequency	150 GHz
Bandwidth per polarization	100 m s^{-1}
Water vapour column density	Default: 5th octile (1.796 mm)
Number of antennas	50 $\times$ 12 m
rms sensitivity	1.5 mJy

speed to within $\sim$250 km s^{-1} (representing an 8 per cent accuracy if the actual speed is 3000 km s^{-1}). This corresponds to determining the redshift difference between the images to 0.1 km s^{-1}.

A flux accuracy of 1.5 mJy corresponds to $\sim$2 per cent accuracy near the peak of the spectral energy distribution. The online calculator suggests that this level of precision can be attained in just over 6 h under typical weather conditions. Thus, we believe that a night with all 50 of the 12 m dishes might allow us to constrain, in this case, the proper motion of the main cluster.

In principle, the subcluster's motion can also be determined using the MCE. However, we could not find known multiple images with separation close to the east–west direction, the likely direction of the collision. Thus, suitable multiple images would first need to be found around the subcluster. This might be accomplished using a fairly deep exposure with the Atacama Pathfinder Experiment (APEX) or other telescopes. If suitable targets were found, they could be targeted for detailed spectroscopic follow-up.

The actual direction of motion of each component of the Bullet Cluster is not known for certain. Thus, observers should target multiple images separated in roughly orthogonal directions. Observing more than one object can also minimize systematics associated with details of the source and the lens.

Once suitable targets are found, we believe that a few nights of observations should be sufficient. The field of view might be large enough that multiple images of different sources can be observed in the same pointing, reducing the required telescope time further. Observing the images simultaneously can also reduce systematics associated with changing atmospheric conditions.

The main difficulty would be in achieving a very accurate calibration of the spectra. However, it is the *relative* redshift between multiple images that is critical for determining the MCE. *Absolute* redshifts are not needed very precisely.

8 CONCLUSIONS

The MCE may provide an essentially direct method to determine the tangential motion of high-z lensing clusters such as the Bullet Cluster, thereby clarifying the tension that appears to exist with ΛCDM (Molnar et al. 2013a). This requires a precise determination of redshift differences between multiple images of the same object.

We expect the MCE to cause multiple images created by the Bullet Cluster to have a redshift velocity difference of $\approx$1 km s^{-1}. We find that, for multiple images of a realistic target, this level of accuracy should be feasible with a night on ALMA, using all 50 of the 12 m dishes. To determine the motions of both the main and the subcluster, multiple pointings may be required.

We considered the effect of the time delay between multiple images. In an expanding Universe, this causes them to have different redshifts (the DEE). However, the effect is second order in the deflection angle, whereas the MCE is first order. Thus, the DEE can be neglected compared with the MCE.

The DME arises when observing an object with a redshift gradient across it, most likely due to rotation. The precise way in which the magnification varies across the source is different for different images. This leads to them having different mean redshifts. Under plausible circumstances, the effect is large enough that it must be considered when trying to infer the lens motion (Fig. 4).

We consider various methods for determining how the DME affects image redshifts. All techniques require the profiles of spectral lines, if only to estimate the redshift gradient across the image based on the linewidth. If the line profile could be observed in more detail, then one could exploit the fact that the DME and MCE affect the line profile in different ways.

Otherwise, the DME could be estimated by determining the parameters which control it (disc scale length and orientation, maximum line of sight rotation speed and sense of rotation, how magnification varies with position in the source plane for each image and, to a smaller extent, the source surface density).

The DME is smaller for some sources than for others. We discuss which types of source might reduce the DME in Section 6. We believe the best option is to use multiple images of an elliptical galaxy as these are likely to rotate slower than spirals, if at all. In particular, the triple image identified in Gonzalez et al. (2009) might be a good source to observe.

The DME is larger for lower mass lenses, making it more important for galaxy–galaxy lensing. Measuring peculiar velocities of galaxies using the MCE might thus be very challenging, especially as these are likely smaller than for the Bullet Cluster.

However, the DME might be easier to observe. This might give more information about the gravitational potential of the lensing galaxy and perhaps the redshift structure of the source. We speculate that the DME might provide a way to estimate the radial gradient in the velocity dispersion of a distant lensed elliptical galaxy.

ACKNOWLEDGEMENTS

The authors wish to thank the referee for several very useful comments. They also thank Keith Horne and Douglas Buisson for helpful discussions. IB is supported by a Science & Technology Facilities Council studentship.

REFERENCES

Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Birkinshaw M., Gull S. F., 1983, Nature, 302, 315
Blaksley C., Bonamente M., 2010, New Astron., 15, 159
Bouillot V. R., Alimi J.-M., Corasaniti P.-S., Rasera Y., 2014, MNRAS, preprint (arXiv:1405.6679)
Cai Y.-C., Cole S., Jenkins A., Frenk C. S., 2010, MNRAS, 407, 201
Cai Y.-C., Li B., Cole S., Frenk C. S., Neyrinck M., 2014, MNRAS, 439, 2978
Clowe D., Gonzalez A., Markevitch M., 2004, ApJ, 604, 596
Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Famaey B., Binney J., 2005, MNRAS, 363, 603
Famaey B., McGaugh S. S., 2012, Living Rev. Relativ., 15, 10
Gómez P. L. et al., 2012, AJ, 144, 79
Gonzalez A. H., Clowe D., Bradač M., Zaritsky D., Jones C., Markevitch M., 2009, ApJ, 691, 525
Ibata R. A. et al., 2013, Nature, 493, 62
Ibata R. A., Ibata N. G., Lewis G. F., Martin N. F., Conn A., Elahi P., Arias V., Fernando N., 2014, ApJ, 784, L6
Jee M. J., Hughes J. P., Menanteau F., Sifón C., Mandelbaum R., Barrientos L. F., Infante L., Ng K. Y., 2014, ApJ, 785, 20
Katz H., McGaugh S., Teuben P., Angus G. W., 2013, ApJ, 772, 10
Kim J., Park C., Gott J. R., III, Dubinski J., 2009, ApJ, 701, 1547
Kovner I., 1990, ApJ, 351, 114
Lage C., Farrar G., 2014, ApJ, 787, 144
Lage C., Farrar G. R., 2015, J. Cosmol. Astropart. Phys., 2, 38
McMillan P. J., 2011, MNRAS, 414, 2446
Mastropietro C., Burkert A., 2008, MNRAS, 389, 967
Menanteau F. et al., 2012, ApJ, 748, 7
Milgrom M., 1983, ApJ, 270, 365
Molnar S. M., Broadhurst T., 2015, ApJ, 800, 37

Molnar S. M., Broadhurst T., Umetsu K., Zitrin A., Rephaeli Y., Shimon M., 2013a, ApJ, 774, 70
Molnar S. M., Chiu I.-N. T., Broadhurst T., Stadel J. G., 2013b, ApJ, 779, 63
Ostriker J. P., Steinhardt P. J., 1995, Nature, 377, 600
Pawlowski M. S., Kroupa P., 2013, MNRAS, 435, 2116
Pawlowski M. S. et al., 2014, MNRAS, 442, 2362
Planck Collaboration XVI, 2014, A&A, 571, A16
Planck Collaboration XXVII, 2014, A&A, 571, A27
Schechter P., 1976, ApJ, 203, 297
Shapiro I. I., 1964, Phys. Rev. Lett., 13, 789
Thompson R., Nagamine K., 2012, MNRAS, 419, 3560
Thompson R., Davé R., Nagamine K., 2014, preprint (arXiv:1410.7438)
Tucker W. H., Tananbaum H., Remillard R. A., 1995, ApJ, 444, 532
Tully R. B., Fisher J. R., 1977, A&A, 54, 661

This paper has been typeset from a TEX/LATEX file prepared by the author.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **459,** 2237–2261 (2016) doi:10.1093/mnras/stw787
Advance Access publication 2016 April 6

Dynamical history of the Local Group in ΛCDM

Indranil Banik★ and Hongsheng Zhao
Scottish Universities Physics Alliance, University of St Andrews, North Haugh, St Andrews, Fife KY16 9SS, UK

Accepted 2016 April 4. Received 2016 March 10; in original form 2015 June 24

ABSTRACT

The positions and velocities of galaxies in the Local Group (LG) measure the gravitational field within it. This is mostly due to the Milky Way (MW) and Andromeda (M31). We constrain their masses using distance and radial velocity (RV) measurements of 32 LG galaxies. To do this, we follow the trajectories of many simulated particles starting on a pure Hubble flow at redshift 9. For each observed galaxy, we obtain a trajectory which today is at the same position. Its final velocity is the model prediction for the velocity of that galaxy. Unlike previous simulations based on spherical symmetry, ours are axisymmetric and include gravity from Centaurus A. We find the total LG mass is $4.33^{+0.37}_{-0.32} \times 10^{12}\,\mathrm{M}_\odot$, with 0.14 ± 0.07 of this being in the MW. We approximately account for IC 342, M81, the Great Attractor and the Large Magellanic Cloud. No plausible set of initial conditions yields a good match to the RVs of our sample of LG galaxies. Observed RVs systematically exceed those predicted by the best-fitting Lambda Cold Dark Matter (ΛCDM) model, with a typical disagreement of $45.1^{+7.0}_{-5.7}$ km s^{-1} and a maximum of $110 \pm 13\,\mathrm{km\,s^{-1}}$ for DDO 99. Interactions between LG dwarf galaxies cannot easily explain this. One possibility is a past close flyby of the MW and M31. This arises in some modified gravity theories but not in ΛCDM. Gravitational slingshot encounters of material in the LG with either of these massive fast-moving galaxies could plausibly explain why some non-satellite LG galaxies are moving away from us even faster than a pure Hubble flow.

Key words: methods: data analysis – methods: numerical – Galaxy: kinematics and dynamics – galaxies: groups: individual: Local Group – cosmological parameters – dark matter.

1 INTRODUCTION

In a homogeneous universe, particles would follow a pure Hubble flow. This means their velocities would depend on their positions according to

$$\boldsymbol{v}(t) = H(t)\,\boldsymbol{r} \quad \text{where } H \equiv \text{Hubble parameter at time } t \qquad (1)$$

However, the Universe is inhomogeneous on small scales. The resulting inhomogeneous gravitational field causes motions to deviate from equation (1). These deviations – termed 'peculiar velocities' – are easily discerned in the Local Group (LG). Thus, the observed positions and velocities of LG galaxies hold important information on the gravitational field in the LG, both now and in the past.[1] Therefore, by investigating a range of physically motivated models for the gravitational field of the LG, we can hope to see which ones – if any – plausibly explain these observations. This technique is known as the timing argument.

The timing argument was first applied to the Milky Way (MW) and Andromeda (M31) galaxies over 50 years ago (Kahn & Woltjer 1959). This pioneering work attempted to match the present relative velocity of the MW and M31, assuming no other major nearby sources of gravity. As M31 must initially have been receding from the MW but is currently approaching it at $\sim$110 km s^{-1} (Slipher 1913; Schmidt 1958), it was clear that models with very little mass in the MW and M31 could not work.[2] In fact, their combined mass had to be $\sim$10 times the observed baryonic mass in these galaxies. This provided one of the earliest indications that most of the mass in typical disc galaxies might be dark.

This conclusion has withstood the test of time, at least in the context of Newtonian gravity. More recent works find a total LG mass of $M \sim (4\text{–}5) \times 10^{12}\,\mathrm{M}_\odot$ (Li & White 2008; van der Marel et al. 2012b; Partridge, Lahav & Hoffman 2013). This is roughly consistent with the combined dynamical masses of the MW and M31. For example, analysis of the giant southern stream around Andromeda (a tidally disrupted satellite galaxy) yielded $M_{\rm M31} \approx 2.5 \times 10^{12}\,\mathrm{M}_\odot$ (Fardal

★ E-mail: ib45@st-andrews.ac.uk

[1] Due to Hubble drag (paragraph below equation 22), peculiar velocities are mostly sensitive to forces acting at late times.

[2] In the limit of no mass, M31 would be receding at $\sim$50 km s^{-1}.

et al. 2013).[3] Combining a wide variety of observations of our own Galaxy, McMillan (2011) found that $M_{\rm MW} \approx 1.5 \times 10^{12}\,{\rm M}_\odot$.[4] However, careful analysis of the Sagittarius tidal stream (Newberg et al. 2002; Majewski et al. 2003) found a mass of about half this (Gibbons, Belokurov & Evans 2014), though this depends on the uncertain distance to the progenitor.

The timing argument seems to suggest a higher mass than the sum of the MW and M31 dynamical masses. The tension would be further exacerbated if the LG mass was smaller in the past, forcing up the present mass inferred by the timing argument. This is quite likely as galaxies accrete mass from their surroundings.

One possible explanation may be that, in the context of a cosmological simulation, the timing argument overestimates the LG mass (González, Kravtsov & Gnedin 2014). However, this trend is not seen in the work of Partridge et al. (2013), whose timing argument calculations included the effect of dark energy. In any case, the tension does not appear to be significant.

The present Galactocentric radial velocity (GRV) of Andromeda provides just one data point. Therefore, it can only be used to constrain one model parameter: the total LG mass. The mass ratio between the MW and M31 cannot be constrained in this way, although it is likely on other grounds that $M_{\rm MW} < M_{\rm M31}$ as M31 is larger (Courteau et al. 2011; Bovy & Rix 2013) and rotates faster (Carignan et al. 2006; Kafle et al. 2012).

More importantly, we cannot determine if the model itself works with just one data point. As a result, it has been suggested to include more distant LG galaxies in a timing argument analysis (Lynden-Bell 1981). Such an analysis was attempted a few years later (Sandage 1986). This work suggested that it was difficult to simultaneously explain all the data then available.

The quality of observational data has improved substantially since that time. More galaxies have also been discovered, providing additional constraints on any model of the LG. This is partly due to wide field surveys such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) and the Pan-Andromeda Archaeological Survey (McConnachie et al. 2009).

Such surveys have shown that satellite galaxies of the MW are preferentially located in a thin (rms thickness ∼25 kpc) corotating planar structure (Pawlowski & Kroupa 2013). Known MW satellites were mostly discovered using the SDSS, which has only limited sky coverage. Even when this is taken into account, it is extremely unlikely that the MW satellite system is isotropic (Pawlowski 2016). In fact, this hypothesis is now ruled out at $>5\sigma$.

A similar pattern is also evident with the satellite galaxies of Andromeda (Ibata et al. 2013). Roughly half of its satellites are consistent with an isotropic distribution but the other half appears to form a corotating planar structure even thinner than that around the MW. However, co-rotation cannot be definitively confirmed until proper motions become available.

The observed degree of anisotropy appears very difficult to reconcile with a quiescent origin in a Lambda-Cold Dark Matter (ΛCDM) universe (Pawlowski et al. 2014, and references therein). This result seems to hold up with more recent higher resolution simulations (Gillet et al. 2015; Pawlowski et al. 2015). One reason is that filamentary infall is unlikely to work because it leads to radial orbits, inconsistent with observed proper motions of several MW satellites (Angus, Diaferio & Kroupa 2011).

[3] This is an estimate of M_{200}.
[4] This is an estimate of the virial mass.

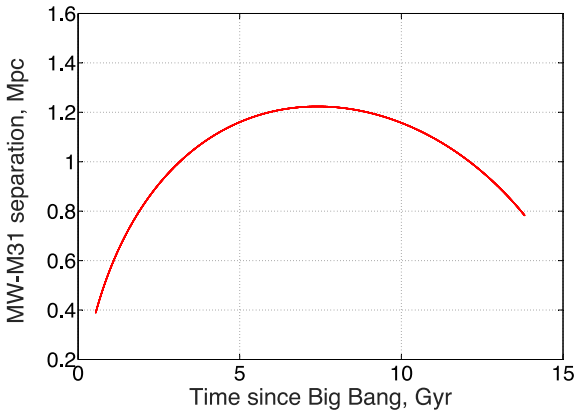


Figure 1. MW–M31 separation d(t) for a typical case where $q_{\rm i} = 0.2$ and $M_{\rm i} = 3.4 \times 10^{12}\,{\rm M}_\odot$. d($t$) always looks broadly similar – in ΛCDM, the MW and M31 have never approached each other closely for any plausible model parameters.

This result has recently been challenged by Sawala et al. (2014) and Sawala et al. (2016) based on the EAGLE simulations, which include baryonic physics (Crain et al. 2015; Schaye et al. 2015). When comparing with the observed satellite systems of the MW and M31, these investigations did not take into account all of the available information, in particular the observed distances to the MW satellites. Once this is considered, it becomes clear that the observed distribution of satellites around the MW is very anisotropic, making a quiescent scenario for their origin much less likely (Pawlowski et al. 2015). Moreover, the inclusion of baryonic physics had very little impact on the extent to which satellite systems are anisotropic. This is what one would expect given the large distances to the MW satellites.

In this context, it seems surprising that a recent investigation found that the observed satellite systems of the MW and M31 are consistent with a quiescent ΛCDM origin at the 5 and 9 per cent levels, respectively (Cautun et al. 2015). However, this analysis suffers from several problems, in particular not considering several objects orbiting the MW (only its 11 classical satellites are considered). The result for the MW is based on assuming that one-third of the sky is not observable due to the Galactic disc. The actual obscured region is likely smaller, making the observed distribution of MW satellites harder to explain. Some of the more important deficiencies with this investigation have been explained by López-Corredoira & Kroupa (2016, last paragraph of page 2).

The MW and M31 are ∼ 0.8 Mpc apart now (McConnachie 2012) and have never interacted in ΛCDM (see Fig. 1). Thus, one might expect their satellite systems to be almost independent in this model. Indeed, it has recently been demonstrated in simulations that the degree of anisotropy of the MW satellite system is not enhanced by the presence of an analogue of M31 (Pawlowski & McGaugh 2014b).

It must be borne in mind that all these authors focused on LG satellites merely because they happen to be nearby, allowing for much more accurate measurements of 3D positions and velocities. It is very difficult to conduct similarly detailed investigations further away. Thus, while it may be dangerous to conclude too much about the Universe based on just ∼ 50 satellite galaxies, one should at least concede that these are located in two essentially independent systems which were not selected because of their anisotropy.

Although a quiescent origin for these highly anisotropic satellite systems appears unlikely, it is possible that an ancient interaction

created them by forming tidal dwarf galaxies (TDGs; Kroupa, Theis & Boily 2005). After all, there are several known cases of galaxies forming from material pulled out of interacting progenitor galaxies (e.g. in the Antennae, Mirabel, Dottori & Lutz 1992).

Such TDGs tend to be more metal rich than primordial galaxies of the same mass (e.g. Croxall et al. 2009). M31 satellites in the planar system around it seem not to have different chemical abundances to M31 satellites outside this plane (Collins et al. 2015). This might be a problem for the scenario, had it involved a *recent* interaction. But with a more ancient interaction, the problem seems to be much less severe (Recchi, Kroupa & Ploeckinger 2015). Essentially, this is because gas in the outer parts of the MW/M31 would have been very metal poor when the interaction occurred. This would lead to TDGs that were initially metal poor, similar to primordial objects of the same age.

TDGs should be free of dark matter as their escape velocity is much below the virial velocity of their progenitor galaxies (Barnes & Hernquist 1992; Wetzstein, Naab & Burkert 2007). Thus, a surprising aspect of LG satellite galaxies is their high mass-to-light (M/L) ratios (e.g. McGaugh & Milgrom 2013). These ratios are calculated assuming dynamical equilibrium. Tides from the host galaxy are probably not strong enough to invalidate this assumption (McGaugh & Wolf 2010). With dark matter unlikely to be present in these systems, the high inferred M/L ratios would need to be explained by modified gravity.

One possibility is to use Modified Newtonian Dynamics (MOND; Milgrom 1983). This imposes an acceleration-dependent modification to the usual Poisson equation of Newtonian gravity (Bekenstein & Milgrom 1984). Despite having only one free parameter, MOND fares well at explaining rotation curves of disc galaxies (Famaey & McGaugh 2012, and references therein). It also seems to work for LG satellites (McGaugh & Wolf 2010; McGaugh & Milgrom 2013), although the relevant observations are challenging.

Applying this theory to the MW and M31, Zhao et al. (2013) found that they would have undergone an ancient close flyby ∼9 billion years (Gyr) ago. The thick disc of the MW would then be a natural outcome of this interaction. Indeed, recent work suggests a tidal origin for the thick disc (Banik 2014). Moreover, its age seems to be consistent with this scenario (Quillen & Garnett 2001).

An ancient flyby of M31 past our Galaxy might have affected the rest of the LG (RLG) as well. Infalling dwarf galaxies might have been flung out at high speeds by gravitational slingshot encounters with the MW/M31. Material might also have been tidally expelled from within them, perhaps forming a dwarf galaxy later on. As a result, the velocity field of the LG would likely have been dynamically heated. We hope to investigate whether there is any evidence for such a scenario.

To this end, the use of more distant LG galaxies can be particularly useful. Within the context of ΛCDM, Peñarrubia et al. (2014) used non-satellite galaxies within ∼ 3 Mpc for a timing argument analysis. Satellite galaxies cannot easily be used in this way because the velocity field becomes complicated close to the MW or M31 (Fig. 3). Intersecting trajectories make it difficult to predict the velocity of a satellite galaxy based solely on its position.

We perform a similar analysis of the same 'target' galaxies as in that work. The basic idea is the same: we construct a test particle trajectory that today is at the same position as a target galaxy.[5] The final velocity relative to the MW is then projected on to our line of sight (equation 50). This model-predicted GRV is corrected for the motion of the Sun with respect to the MW, yielding a heliocentric radial velocity (HRV) prediction which can be compared with observations. When proper motion measurements become available, it will be very interesting to compare the full 3D velocities of LG galaxies with our models.

For simplicity, we assume that the only massive objects in the LG are the MW and M31, which we take to be on a radial orbit. Recent proper motion measurements of M31 indicate only a small tangential motion relative to the MW (van der Marel et al. 2012a). This makes the true orbit almost radial.

The recent work of Salomon et al. (2016) argues for a high M31 proper motion ($\sim 100\,\mathrm{km\,s^{-1}}$) based on redshift gradients in the M31 satellite system. This measurement is consistent with the more direct measurement of van der Marel et al. (2012a), though there is some tension. This might be explained by intrinsic rotation of the M31 satellite system. With a field of view of perhaps 5°, rotation at only an $\sim 10\,\mathrm{km\,s^{-1}}$ level can masquerade as a proper motion of $\sim 100\,\mathrm{km\,s^{-1}}$. In fact, there is strong evidence that nearly half of the M31 satellites rotate coherently around it (Ibata et al. 2013). Although Salomon et al. (2016) take this into account to some extent, other rotating satellite planes might also exist around M31. This is suggested by recent investigations into the kinematics of its globular cluster system (Veljanoski et al. 2014). Moreover, a large tangential velocity between the MW and M31 would show up as a dipole-like feature in the radial velocities of distant LG galaxies. This has been searched for but not found (Peñarrubia et al. 2016). Thus, we assume that the van der Marel et al. (2012a) proper motion measurement is more accurate, making the MW–M31 orbit nearly radial.

Starting at some early initial time t_i, we evolved forwards a large number of test particles in the gravitational field of the MW and M31. We took the barycentre of the LG at $t = t_i$ as the centre of the expansion. The initial velocities followed a pure Hubble flow (equation 37). This is because the Universe was nearly homogeneous at early times – peculiar velocities on the last scattering surface are only $\sim 1\,\mathrm{km\,s^{-1}}$ (Planck Collaboration XIII 2015), much less than typical values today ($\sim 50\,\mathrm{km\,s^{-1}}$, see Fig. 11).

As both the initial conditions and the gravitational field are axisymmetric, test particles move within meridional planes (i.e. those containing the symmetry axis). This allowed us to use an axisymmetric model. We briefly mention that the gravitational field in our model varies with time, because the MW and M31 move.

A major improvement of our analysis is that the LG is not treated as spherically symmetric. This assumption is not a very good one as the targets considered are at distances of ∼1–3 Mpc. Meanwhile, the MW–M31 separation is ∼0.8 Mpc (McConnachie 2012). This means that the gravitational potential – and thus velocities – are likely to deviate substantially from spherical symmetry in the region of interest. However, we expect only small deviations from axisymmetry for reasons just stated.

Objects outside the LG can have some influence on our results because they can raise tides on the LG. The most important perturbers were identified by Peñarrubia et al. (2014) as M81, IC 342 and Centaurus A. Their properties are given in Table 1. We directly included the gravity of the most massive of these objects, Cen A (Section 2.2.2). We took advantage of its location on the sky being almost exactly opposite that of Andromeda. Due to the large distance of Cen A from the LG (∼ 4 Mpc), its velocity is dominated by the Hubble expansion (Karachentsev et al. 2007). This makes the LG–Cen A trajectory almost radial, allowing us to continue using our axisymmetric model.

[5] Our model is effectively two-dimensional, so we used a 2D version of the Newton–Raphson algorithm to achieve this.

Table 1. Properties of mass concentrations outside the LG which we considered. Distances are in Mpc and sky positions are in the Galactic system. The estimate for Cen A is from Harris, Rejkuba & Harris (2010), that of M81 is from Gerke et al. (2011) and that of IC 342 is from Wu et al. (2014). Masses are in units of $10^{12}\,M_\odot$ and were obtained from Karachentsev (2005) for Cen A and IC 342. For M81, we used Karachentsev & Kashibadze (2006).

Name	b	l	$d_{\rm MW}$	M
Centaurus A	19.°4173	309.°5159	3.8	4
M81	40.°9001	142.°0918	3.6	1.03
IC 342	10.°5799	138.°1726	3.45	1.76

Our paper is structured as follows: the governing equations and methods are described in Section 2. This section also shows some results, to give a rough idea of what happens in our simulations. Comparison of simulation outputs with observations is done in Section 3. The posterior probability density functions of all variables and pairs of variables are shown in Fig. 7. Our results indicate that no model comes close to reproducing all the observations simultaneously.

In Section 4, we discuss several shortcomings of our model and whether accounting for some of them might help to explain the observations. In Table 5, we show how Cen A affects our results. We also estimate carefully the effects of M81 and IC 342 (Table 6), the Great Attractor (GA, Fig. 16) and the Large Magellanic Cloud (LMC, Fig. 17). These objects seem to little affect GRVs and often worsen the discrepancy with the best-fitting model. We suggest a possible explanation for our results in Section 4.6. Differences between our approach and the similar study of Peñarrubia et al. (2014) are described in Section 4.7. Our conclusions are summarized in Section 5.

2 METHOD

The method we follow is to ensure a simulated test particle ends up at the same position as each LG galaxy in our sample (a 'target'). At present, only the radial velocities of our targets are available. Thus, the velocity of this particle relative to that of the MW is projected on to the direction towards the particle (equation 50). This model-predicted GRV is then corrected for solar motion in the MW and compared with observations. The procedure is repeated for different model parameters, which are systematically varied across a grid. Therefore, within the priors we set (Table 2), all model parameter combinations were investigated.

2.1 Equations of motion

We begin with the metric in the weak field limit

$$\mathrm{d}s^2 = c^2\mathrm{d}\tau^2 \tag{2}$$

$$= \left(1+\frac{2\Phi}{c^2}\right)c^2\mathrm{d}t^2 - \left(1-\frac{2\Phi}{c^2}\right)a^2\left(\mathrm{d}\chi^2 + S^2(\chi)\,\mathrm{d}\Omega^2\right) \tag{3}$$

$$S(\chi) \equiv \begin{cases} \sinh(\chi) & \text{in an open universe} \\ \chi & \text{in a flat universe} \end{cases} \tag{4}$$

$$C(\chi) \equiv \begin{cases} \cosh(\chi) & \text{in an open universe} \\ 1 & \text{in a flat universe}\,. \end{cases} \tag{5}$$

Table 2. Priors and 1σ confidence levels on model parameters. The latter are far from the boundaries imposed by the former, showing that our results are not strongly affected by our priors. Due to accretion, the present-day masses of the MW and M31 are ∼5 per cent higher than when the simulations start. We use the measurement of d_0 by McConnachie (2012). Cosmological parameters are from Planck Collaboration XIII (2015). We obtained $v_{c,\odot}$ from McMillan (2011) and the Sun's non-circular velocity from Francis & Anderson (2014). Uncertainty in the latter is much less than in the former. We assume $v_{c,\odot}$ is within 3σ of its most likely value.

Name	Meaning and units	Prior	Result
$\sigma_{\rm extra}$	Extra velocity dispersion along line of sight, km s^{-1}	0–100	$45.1^{+7.0}_{-5.7}$
M_i	Initial MW + M31 mass, trillions of solar masses	2–6.6	4.1 ± 0.3
q_1	Fraction of MW + M31 mass initially in the MW	0.04–0.96	0.14 ± 0.07
$v_{c,\odot}$	Circular speed of MW at position of Sun, km s^{-1}	239 ± 5	239.5 ± 4.8
Fixed parameters			
d_0	Distance to M31, kpc	783 ± 25	
H_0	Hubble constant at the present time, km s^{-1} Mpc^{-1}	67.3	
$\Omega_{\rm m,0}$	Present matter density in the Universe $\div \frac{3H_0^2}{8\pi G}$	0.315	
a_i	Scalefactor of Universe at start of simulation	0.1	
$r_{\rm acc,MW}$	Accretion radius of MW	15 337 parsecs	
$r_{\rm acc,M31}$	Accretion radius of M31	21 472 parsecs	
$U_\odot$	See equation (47)	14.1 km s^{-1}	
$V_\odot$	See equation (47)	14.6 km s^{-1}	
$W_\odot$	See equation (47)	6.9 km s^{-1}	

Here, c is the speed of light and τ is proper time. The scalefactor of the Universe is a. The spatial part of the metric has been written in spherical polar coordinates, with $\mathrm{d}\Omega$ representing a change in angle. Using the coordinates $x^0 \equiv t$, $x^1 \equiv \chi$, $x^2 \equiv \theta$ and assuming spherical symmetry, we get a diagonal metric where

$$g_{00} = c^2\left(1+\frac{2\Phi}{c^2}\right) \tag{6}$$

$$g_{11} = -a^2\left(1-\frac{2\Phi}{c^2}\right) \tag{7}$$

$$g_{22} = -a^2\left(1-\frac{2\Phi}{c^2}\right)S^2(\chi) = g_{11}S^2(\chi)\,. \tag{8}$$

As the metric coefficients are independent of $x^2 \equiv \theta$, the geodesic equation tells us that

$$\dot{x}_2 = \sum_{b=0}^{3} g_{2b}\dot{x}^b \quad (\text{only non}-\text{zero term is } b=2) \tag{9}$$

$$= a^2\left(1-\frac{2\Phi}{c^2}\right)S^2(\chi)\,\dot{\theta}$$

$$= \text{constant}. \tag{10}$$

We use $\dot{q}$ to denote the derivative of any quantity q with respect to proper time. In weak gravitational fields ($\Phi \ll c^2$), proper and coordinate time are almost equal, making $\tau \approx t$. Bearing this in mind, equation (10) tells us that the specific angular momentum of a test particle is conserved. This is due to the spherical symmetry of the situation.

For the radial component of the motion, we use the geodesic equation in the form

$$\ddot{x}^1 + \sum_{b=0}^{3}\sum_{c=0}^{3} \Gamma^1{}_{bc}\dot{x}^b\dot{x}^c = 0 \text{ where} \tag{11}$$

$$\Gamma^a{}_{bc} = \frac{1}{2}\sum_{d=0}^{3} g^{ad}\left(\partial_b g_{dc} + \partial_c g_{bd} + \partial_d g_{bc}\right). \tag{12}$$

Here, we use the notation $\partial_b q \equiv \frac{\partial q}{\partial x^b}$ for any quantity q. The non-zero Christoffel symbols relevant to a non-relativistic test particle in this situation are

$$\Gamma^1{}_{00} \approx \frac{\Phi'}{a^2} \tag{13}$$

$$\Gamma^1{}_{01} \approx \frac{\dot{a}}{a} \equiv H \tag{14}$$

$$\Gamma^1{}_{22} \approx -S(\chi)C(\chi) \tag{15}$$

$$\Gamma^1{}_{11} \approx -\frac{\Phi'}{c^2}. \tag{16}$$

Here, q' implies a partial derivative with respect to the comoving coordinate χ rather than physical distance $a\chi$. Putting in the non-negligible Christoffel symbols[6] into equation (11), we get that

$$\ddot{\chi} + \frac{\Phi'}{a^2} + 2H\dot{\chi} - S(\chi)C(\chi)\dot{\theta}^2 = 0. \tag{17}$$

In terms of physical coordinates $r \equiv a\chi$, this becomes

$$\ddot{r} = \left(\frac{\ddot{a}}{a}\chi + 2H\dot{\chi} + \ddot{\chi}\right)a \tag{18}$$

$$= \left(\frac{\ddot{a}}{a}\chi - \frac{1}{a^2}\frac{\partial\Phi}{\partial\chi} + S(\chi)C(\chi)\dot{\theta}^2\right)a \tag{19}$$

$$= \frac{\ddot{a}}{a}r - \frac{\partial\Phi}{\partial r} + S(\chi)C(\chi)a\dot{\theta}^2. \tag{20}$$

The real Universe is close to spatially flat (Planck Collaboration XIII 2015). Thus, from now on, we will only consider the case of a flat universe. This is defined as one having a density ρ equal to the critical density $\rho_{\rm crit}$, if we count both matter and dark energy towards the total density.

$$\rho = \rho_{\rm crit} \equiv \frac{3H^2}{8\pi G}. \tag{21}$$

Equation (21) is valid at all times, although both ρ and H vary with time. In such a universe, equation (20) becomes

$$\ddot{r} = \frac{\ddot{a}}{a}r - \frac{\partial\Phi}{\partial r} + r\dot{\theta}^2. \tag{22}$$

This looks very similar to the Newtonian equation of motion. The last term corresponds to the centrifugal force, while the $\frac{\partial\Phi}{\partial r}$ term corresponds to the potential gradient. The only novel aspect is the term $\frac{\ddot{a}}{a}r$. The importance of this term becomes clear if we consider a homogeneous universe, meaning that $\Phi = 0$ everywhere and at all times. In this case, we expect the distance between two non-interacting test particles to behave as $r \propto a(t)$ (i.e. their comoving distance is constant). This implies that $\ddot{r} = \frac{\ddot{a}}{a}r$. This term has also been called Hubble drag because it tends to reduce the magnitude of peculiar velocities.[7] In the absence of potential gradients, we would get $v_{pec} \propto \frac{1}{a(t)}$, where the peculiar velocity is defined by

$$\boldsymbol{v}_{\rm pec} \equiv \dot{\boldsymbol{r}} - H\boldsymbol{r}. \tag{23}$$

In general, the Universe is neither homogeneous nor spherically symmetric. For such circumstances, we suppose that the generalization of equation (22) is given by

$$\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r} - \nabla\Phi. \tag{24}$$

With the equations of motion in hand, we now need to relate the potential Φ to the density perturbations that act as its source. To do this, we use the 00 component of the field equation of General Relativity.

$$R_{ab} = -\frac{\pi}{c^4}\left(T_{ab} - \frac{1}{2}Tg_{ab}\right) \text{ where} \tag{25}$$

$$T \equiv \sum_{d=0}^{3} T^d{}_d. \tag{26}$$

Here, T_{ab} is the energy-momentum tensor while R_{ab} is the Ricci tensor, related to the curvature of the metric. Perturbations to the solution for a homogeneous universe must satisfy the equation

$$\Delta R_{00} = -\frac{8\pi G}{c^4}\Delta\left(T_{00} - \frac{1}{2}Tg_{00}\right). \tag{27}$$

In this case, for non-relativistic sources which are almost pressureless (like baryons and cold dark matter), we get that

$$R_{00} \approx -\sum_{i=1}^{3}\frac{\partial_i\left(\partial_i\Phi\right)}{a^2c^2} \text{ (sum over spatial indices only).} \tag{28}$$

The stress-energy tensor takes on a particularly simple form: its only non-zero element is $T_{00} \approx \rho c^2$. Thus,

$$g_{00}T \approx T_{00} \approx \rho c^2. \tag{29}$$

Using equations (28) and (29) in equation (25), we get that

$$\nabla^2\Phi = 4\pi G\Delta\rho. \tag{30}$$

Here, $\nabla^2\Phi$ is the Laplacian of Φ with respect to physical coordinates. In spherical symmetry, it is

$$\nabla^2\Phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right). \tag{31}$$

Equation (30) is very similar to the usual Poisson equation of Newtonian gravity. Note, however, that only deviations from the background density act as a source for Φ (i.e. it is sourced by $\Delta\rho$ rather than ρ).

2.2 Simulations

2.2.1 Including the MW and Andromeda

The LG is assumed to consist of two point masses (the MW and M31) plus a uniform distribution of matter at the same density as the

[6] The $\Gamma^1{}_{11}$ term effectively causes $\Phi' \to \Phi'\left(1 - \frac{v^2}{c^2}\right)$, where v is the speed of the particle with respect to a comoving observer at the same place. This leads to a special relativistic correction which makes it difficult for a potential gradient to accelerate a particle if its speed is close to that of light. For non-relativistic particles, the effect of this term is negligible because $v \ll c$.

[7] If the Universe were contracting, then this term would be a forcing to peculiar velocities rather than a drag upon them.

cosmic mean value $\overline{\rho}$ (see Section 4.1 for further discussion of this point). Our simulations start when the scalefactor of the Universe $a = a_i$. We used $a_i = 0.1$, though our results change negligibly if $a_i = \frac{1}{15}$ instead (see Section 4.6).

The initial separation of the MW and M31 d_i is varied to match their presently observed separation d_0 using a Newton–Raphson technique. Note that altering d_i alters their initial velocities because the galaxies are assumed to have zero peculiar velocity at the start of the simulation ($t = t_i$). Thus, their final attained separation d_f depends strongly on d_i.

The MW–M31 orbit is taken to be radial, a reasonable assumption given their small tangential motion ($\sim$17 km s^{-1} compared to a radial velocity of $\sim$110 km s^{-1}; van der Marel et al. 2012a). This makes the gravitational field in the LG axisymmetric. As the initial conditions are spherically symmetric (equation 37), a 2D model is sufficient for this investigation.

Applying equation (24) to a radial orbit, the distance d between the galaxies satisfies

$$\ddot{d} = -\frac{GM}{d^2} + \frac{\ddot{a}}{a} d \tag{32}$$

$$\dot{d} = H_i d_i \quad \text{initially.} \tag{33}$$

H_i is the value of the Hubble constant $\frac{\dot{a}}{a}$ when $t = t_i$, while d_i is the MW–M31 separation at that time. M is the combined mass of the MW and M31. It can be verified straightforwardly that when $M = 0$ (i.e. non-interacting test particles), we recover $d \propto a$. In this case, the galaxies trace the cosmic expansion but do not influence each other.

Equation (32) implicitly assumes that the MW and M31 are surrounded by a distribution of matter with the same density as the cosmic mean value $\overline{\rho}$. This point is discussed more thoroughly in Section 4.1, where we also redo our entire analysis assuming instead that the surroundings of the MW and M31 are empty.

We use a standard flat[8] dark energy dominated cosmology with parameters given in Table 2. Therefore,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho_m - 2\rho_\Lambda) \tag{34}$$

$$= {H_0}^2 \left(-\frac{1}{2}\Omega_{m,0}\, a^{-3} + \Omega_{\Lambda,0}\right). \tag{35}$$

Defining time t to start when $a = 0$ and requiring that $\dot{a} = H_0$ when $a = 1$ (the present time), we get that

$$a(t) = \left(\frac{\Omega_{\rm m,0}}{\Omega_{\Lambda,0}}\right)^{1/3} \sinh^{2/3}\left(\frac{3}{2}\sqrt{\Omega_{\Lambda,0}} H_0 t\right). \tag{36}$$

The present values of H_0 and $\Omega_{\rm m,\,0}$ uniquely determine the present age of the Universe t_f via inversion of equation (36) to solve for when $a = 1$. We also use it to determine when $a = a_i$, thereby fixing the start time of our simulations.

The timing argument is particularly sensitive to late times (Fig. 4). This makes it important to correctly account for the late-time effect of dark energy. Because this tends to increase radial velocities of LG galaxies, one is forced to increase the mass of the LG to bring their predicted radial velocities back down to the observed values. As a result, the inclusion of dark energy in timing argument analyses of the LG increases its inferred mass by a non-negligible amount (Partridge et al. 2013).

[8] $\Omega_{\rm m,\,0} + \Omega_{\Lambda,\,0} = 1$

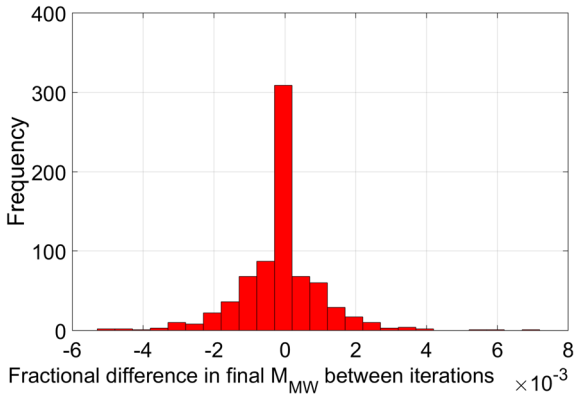


Figure 2. Fractional difference between the final attained MW mass on the first and second runs of each simulation. The very small values show that our solution for $M_{\rm MW}(t)$ converged well. Simulations without Centaurus A give similar results.

Once we obtained a trajectory that (very nearly) satisfied $d_f \equiv d(t_f) = d_0$, we had the ability to find the gravitational field everywhere in the LG at all times. A large number of test particles were then evolved forwards, all starting on a pure Hubble flow with the centre of expansion at the barycentre of the LG.

$$\mathbf{v}_i = H_i \boldsymbol{r}_i. \tag{37}$$

Note that equation (37) also applies to the MW and M31, which we model as point masses. A point mass approximation should work for determining d(t) as Andromeda never gets very close to the MW (Fig. 1). However, it is not good for handling close encounters of test particles with either galaxy. Thus, we adjust the forces they exert on test particles to be $\propto \frac{1}{r}$ at low r (i.e. close to the attracting body). This is for consistency with the observed flat rotation curves of the MW and M31. To recover $g \propto \frac{1}{r^2}$ at large r, we set the gravity towards each galaxy to be

$$g = \frac{GM}{r^2}\frac{b}{\sqrt{1+b^2}} \quad \text{where} \tag{38}$$

$$b \equiv \frac{r}{r_s}. \tag{39}$$

r_s is chosen so that the force at $r \ll r_s$ leads to the correct flatline level of rotation curve for each galaxy, i.e. $r_s = \frac{GM}{{v_f}^2}$. For the MW, we take $v_f = 180$ km s^{-1} (Kafle et al. 2012) while for Andromeda, we take $v_f = 225$ km s^{-1} (Carignan et al. 2006).

Combining equation (38) with the cosmological acceleration term, the equation of motion for our test particles is

$$\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r} - \sum_{j={\rm MW,M31}} \frac{GM_j\left(\boldsymbol{r} - \boldsymbol{r}_j\right)}{\left(|\boldsymbol{r} - \boldsymbol{r}_j|^2 + {r_{s,j}}^2\right)^{1/2}|\boldsymbol{r} - \boldsymbol{r}_j|^2}. \tag{40}$$

Some trajectories go very close to the MW or M31. Approaches within a distance of $r_{\rm acc}$ (given in Table 2) are handled by terminating the trajectory and assuming the particle was accreted by the nearby galaxy. This causes the mass of that galaxy to increase.

As we solved the test particle trajectories sequentially, it was not possible until the very end to have the mass histories $M_{\rm MW}(t)$ and $M_{\rm M31}(t)$. Thus, we assumed constant masses for the force calculations. We then repeated the process, using the previously stored mass histories for each galaxy. This meant that the initial MW–M31 separation d_i also had to be adjusted. In this way, we found that the final mass had converged fairly well with just two iterations (Fig. 2).

The changing mass of the MW and M31 meant that one could not trivially convert the separation history d(t) into MW and M31 positions ($y_{\rm MW}(t)$ and $y_{\rm M31}(t)$, respectively). However, the instantaneous acceleration of the MW must be due to the gravity of Andromeda.[9] Thus, the magnitude of this acceleration must be a fraction $\frac{M_{\rm M31}}{M_{\rm MW}+M_{\rm M31}}$ of the total mutual acceleration. This means that

$$\ddot{y}_{\rm MW}(t) = \frac{M_{\rm M31}(t)}{M_{\rm MW}(t)+M_{\rm M31}(t)}\,\ddot{d} \tag{41}$$

$$\ddot{y}_{\rm M31}(t) = -\frac{M_{\rm MW}(t)}{M_{\rm MW}(t)+M_{\rm M31}(t)}\,\ddot{d}. \tag{42}$$

In practice, we solved equation (32) to determine d(t). We found the change in separation over each time timestep d($t+\Delta t$) − d(t) and apportioned this to the MW and M31 in inverse proportion to their masses at $t+\frac{1}{2}\Delta t$.

Our equations are referred to the frame of reference in which the origin corresponds to the initial centre of mass position (considering only the MW and M31). This makes our reference frame inertial. We do not keep track of how the centre of mass moves after our simulations start.

The initial masses of the MW and M31 imply that they must have accreted material in some region prior to the start of our simulation. Thus, we do not allow test particles to start within a certain excluded region. This is defined by an equipotential $U_{\rm exc}$, chosen so as to enclose the correct total volume (i.e. $V_{\rm exc}\,\rho_{M,i} = M_i$, the initial LG mass). The density of matter $\rho_{M,i}$ at the initial time t_i includes contributions from both baryonic and dark matter. For most parameters, the resulting excluded region is a single region encompassing both the MW and M31 rather than distinct regions around each galaxy.

The potential resulting from integrating equation (38) is

$$U = \sum_{j={\rm MW,M31}} \frac{GM}{r_{S,j}} \ln\left(\frac{\sqrt{1+{b_j}^2}-1}{b_j}\right). \tag{43}$$

We start our test particles on a grid of plane polar coordinates. At some particular angle θ, we consider a sequence of trajectories which start further and further out. Trajectories are skipped if they start within the 'exclusion zone' ($U < U_{\rm exc}$ at $t = t_i$). Once we obtain a trajectory that finishes further than 2.15 Mpc from the LG barycentre, we skip three out of every four steps as the velocity field is fairly smooth at such large distances (Fig. 3). Once we reach beyond 3.2 Mpc, we move on to the next value of θ. This is because we do not need the velocity field further than ∼3 Mpc from the LG as there are no target galaxies further away.[10]

We use a fourth-order Runge–Kutta algorithm with an adaptive timestep designed to be 30–70 times shorter than the instantaneous dynamical time $t_{\rm dyn}$. This is estimated by dividing the distance to each galaxy by the speed of a test particle with zero total energy, ignoring the presence of the other galaxy. Faced with two estimates of $t_{\rm dyn}$, we use the shorter one in order to maximize the resolution.

[9] This is not strictly true at early times due to the $\frac{\ddot{a}}{a}y$ term, but we do not expect either galaxy to have accreted much mass at that stage because no test particle starts very close to the MW or M31. Without mass accretion, the ratio of this term between the two galaxies is also inverse to that of their masses.

[10] This requires trajectories starting out to distances of ∼0.5 Mpc.

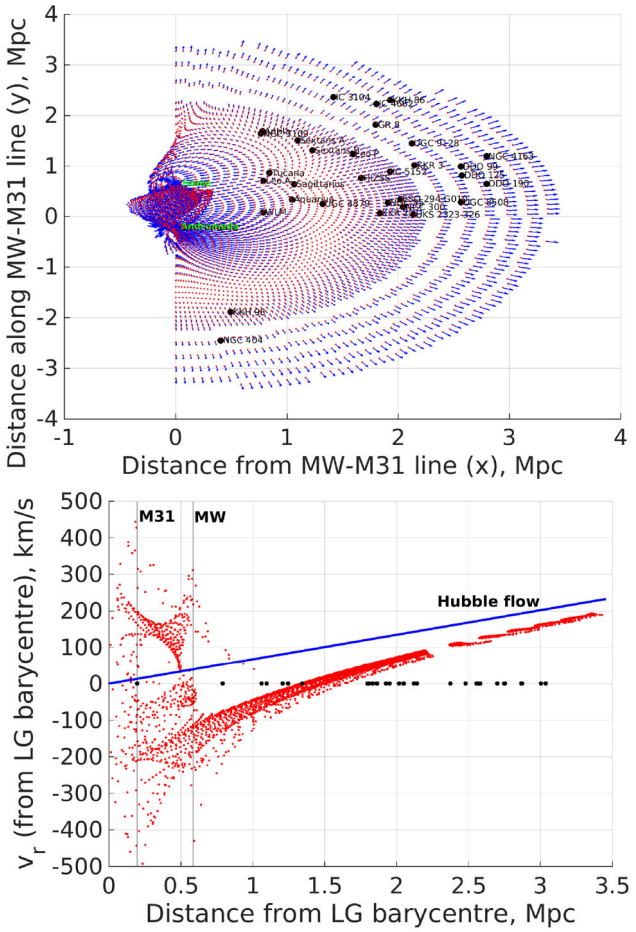


Figure 3. Top: LG velocity field for the case $q_1 = 0.3$, $M_i = 4\times10^{12}\,{\rm M}_\odot$ and no Centaurus A. Locations of indicated galaxies are shown. The MW is just above the origin. Only particles starting at $x \geq 0$ (and thus $v_x \geq 0$) were considered. Thus, the presence of particles at $x < 0$ indicates intersecting trajectories and a disturbed velocity field. Bottom: radial velocities of test particles with respect to the LG barycentre. Vertical lines represent the distances of M31 and the MW from there. Increased velocity dispersion near these galaxies is apparent. Black dots on the x-axis show the distances of target galaxies from the LG barycentre. Without proper motions, observations cannot be put on such a Hubble diagram as the MW is not at the LG barycentre.

The worst time resolution we use is $\frac{1}{1000}$ of the total duration (∼13.5 Gyr). This was sufficient for distances $\gtrsim r_s$ from each galaxy. At smaller distances, we found that $t_{\rm dyn} \propto r^{5/4}$ is a good approximation. If required, we improve the time resolution in powers of 2 up to a maximum of 5 times (for a $2^5 = 32\times$ reduction in Δt). This should provide adequate resolution for distances from the MW and M31 greater than their respective 'accretion radii' $r_{\rm acc}$, which we chose to be a few disc scalelengths (Table 2).

2.2.2 *Including Centaurus A*

Although none of our target galaxies are too close to any of the perturbers listed in Table 1 (due to pre-selection by Peñarrubia et al. 2014), we were still concerned that their gravity might have noticeably affected our target galaxies. To test this scenario, we decided to directly include the most massive perturber, Centaurus A.

Due to the large distance of Cen A from the LG (∼4Mpc), any peculiar velocity it has is likely to be much smaller than its radial velocity. Indeed, this is borne out observationally for motion along our line of sight (Karachentsev et al. 2007). As a result, Cen A is probably on an almost radial orbit with respect to the LG barycentre. Fortunately, Cen A is currently located almost directly opposite M31 on our sky ($\cos\theta = -0.99$, where θ is the angle on the sky between M31 and Cen A). This allowed us to continue using our axisymmetric model.

To initialize each simulation, we need trajectories for the MW, M31 and Cen A that match the presently observed distances to M31 and Cen A. This is done using a 2D Newton–Raphson algorithm[11] on the initial relative positions of all three galaxies along a line. As before, initial velocities were found using equation (37).

Test particle trajectories were then solved in the usual way, with the grid of initial positions centred on the initial barycentre of the MW and M31 as before. Including Cen A, equation (40) becomes

$$\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r} - \sum_{\substack{j=\mathrm{MW},\\ \mathrm{M31,CenA}}} \frac{GM_j\left(\boldsymbol{r}-\boldsymbol{r}_j\right)}{\left(|\boldsymbol{r}-\boldsymbol{r}_j|^2 + {r_{S,j}}^2\right)^{1/2}|\boldsymbol{r}-\boldsymbol{r}_j|^2}. \tag{44}$$

For simplicity, we keep the mass of Cen A fixed at $4\times10^{12}\,\mathrm{M}_\odot$ (Karachentsev 2005) but still allow the MW and M31 to accrete mass.

2.3 Observations and sample selection

Our data set comes mostly from the catalogue of LG galaxies compiled by McConnachie (2012). We used the subset of these that were used for a timing argument analysis by Peñarrubia et al. (2014). This implicitly applies a number of criteria. The basic idea behind them was to ensure that gravity from the MW and M31 dominates over gravity from anything else. For this reason, targets $\gtrsim$3 Mpc from the LG were not considered. The authors also avoided galaxies too close to any major mass concentrations outside the LG. The perturbers they considered are listed in Table 1.

Very close to the MW or M31, there are crossing trajectories and so the model-predicted velocity in such regions is not well defined (top panel of Fig. 3). Further away, this issue does not arise. Thus, Peñarrubia et al. (2014) restricted their sample to *non-satellite* galaxies. We further restricted their sample by excluding Andromeda XVIII as it is in the disturbed region around M31, even if it is unbound.

We treated HIZSS 3A and B as one object as they are almost certainly a binary system. Naturally, we used the velocity of its centre of mass, assuming a mass ratio of 13:1 (Begum et al. 2005). To allow for uncertainty in this ratio, we inflated the error on the HRV to $3.5\,\mathrm{km\,s^{-1}}$. This decision turns out not to matter very much because the uncertainty in its distance has a much larger effect than uncertainty in its HRV (this is true for most of our targets – distances are harder to measure accurately).

In regions close to the MW and M31, the presence of crossing trajectories makes it impossible to uniquely predict the velocity of a target galaxy based solely on its position (Fig. 3). In such cases, we should reject the target (i.e. not use it in our analysis). In practice, we accepted all of our targets in all cases. We checked the velocity field to ensure none of our target galaxies fell in regions with crossing trajectories. Although none of them did so, NGC 3109 and Antila came close. We tried raising H_0 and altering the distances to these galaxies within their uncertainties, but we still could not get them in a region of crossing trajectories. In any case, excluding them would not much affect our conclusions, as will become apparent later (Fig. 15). As a final check, IB looked at all the $\sigma_{\rm pos}$ and GRV maps (like those in Fig. 5) and confirmed that they were smooth.

If we had been less fortunate regarding the locations of our target galaxies, then we might have rejected some of them in some simulations using criteria designed to search for intersecting trajectories. The best options seem to be a high density of test particles near the present position of the target and a high velocity dispersion at that position. In this case, we might have to alter equation (57) by multiplying the first term on the right by $\frac{32}{N}$, where N is the number of target galaxies 'accepted' by the algorithm. Additional care would have to be taken to ensure the analysis remained valid despite N varying with the model parameters (i.e. some models might be constrained using fewer observations than others).[12]

To convert observations into the same coordinates as our simulations, we first defined Cartesian xy coordinates centred on the LG barycentre, with $\hat{\boldsymbol{y}}$ towards the MW. The positions of observed galaxies were converted into this system using the equations

$$x = d_{\rm MW}\,|\hat{\boldsymbol{d}}_{\rm MW}\times\hat{\boldsymbol{r}}_{\rm MW}| \tag{45}$$

$$y_{\rm rel} \equiv y - y_{\rm MW} = d_{\rm MW}\left(\hat{\boldsymbol{d}}_{\rm MW}\cdot\hat{\boldsymbol{r}}_{\rm MW}\right), \tag{46}$$

where $y_{\rm MW}$ is the present distance of the MW from the initial position of the LG barycentre. $\hat{\boldsymbol{y}}\equiv\hat{\boldsymbol{r}}_{\rm MW}$ is the direction from M31 towards the MW. This is just the opposite of the direction in which we observe Andromeda. $d_{\rm MW}$ is the distance from the MW to the target galaxy. This is essentially equivalent to its heliocentric distance. We neglected the difference that arises because the Sun is not at the centre of the MW.[13] For this reason, we can approximate the direction between the MW centre and the target galaxy $\hat{\boldsymbol{d}}_{\rm MW}$ as the direction in which we observe it.

Although the position of the Sun with respect to the Galactic Centre is unimportant for this work, its *velocity* relative to the MW is very important because this velocity is $\sim250\,\mathrm{km\,s^{-1}}$ (McMillan 2011). For observational reasons, we split this velocity into two components. The MW is a disc galaxy, so most of the Sun's velocity is just ordered circular motion within the disc plane. In the absence of non-circular motions, its speed would be $v_{c,\odot}$. This is known as the local standard of rest (LSR) because particles moving tangentially at this speed would be at rest in a rotating reference frame.

We temporarily define a 3D Cartesian coordinate system with $\hat{\boldsymbol{x}}$ pointing from the Sun towards the Galactic Centre, $\hat{\boldsymbol{z}}$ pointing towards the North Galactic Pole and $\hat{\boldsymbol{y}}$ chosen so as to make the system right handed. Fortunately, $\hat{\boldsymbol{y}}$ points along the direction of rotation. In this system, the velocity of the Sun with respect to the

[11] For stability, we underrelaxed the algorithm, meaning that in each iteration, we altered the parameters by 80 per cent of what the algorithm would normally have altered them by.

[12] If a target galaxy is problematic in only some parts of parameter space, then one can simply avoid including it in the analysis altogether, thereby avoiding issues due to N being model dependent. However, this makes poorer use of the available information.

[13] Target galaxies are $\gtrsim$800 kpc away while the Sun is only ∼8kpc from the Galactic Centre, well below typical distance errors.

MW (including its non-circular motion) is

$$\boldsymbol{v}_\odot = \begin{bmatrix} U_\odot \\ V_\odot + v_{c,\odot} \\ W_\odot \end{bmatrix}. \tag{47}$$

The direction towards another galaxy can be determined from its Galactic coordinates using

$$\hat{\boldsymbol{d}}_{\rm MW} = \begin{bmatrix} \cos b\ \cos l \\ \cos b\ \sin l \\ \sin b \end{bmatrix}, \tag{48}$$

where b is the Galactic latitude and l is the Galactic longitude, whose zero-point is the direction towards the Galactic Centre. Galactic coordinates are actually heliocentric, though the distinction is unimportant for very distant objects.

Without proper motions of LG galaxies, only their GRV can be constrained. Thus, we project the velocity of the Sun with respect to the MW on to the direction towards the desired galaxy. This is then added on to its observed HRV.

$$\mathrm{GRV}_{\rm obs} = \mathrm{HRV}_{\rm obs} + \boldsymbol{v}_\odot \cdot \hat{\boldsymbol{d}}_{\rm MW}. \tag{49}$$

This estimate of $\mathrm{GRV}_{\rm obs}$ is dependent on the model used for $\boldsymbol{v}_\odot$, in particular the adopted LSR speed. Thus, for a range of plausible values of $v_{c,\odot}$, we stored the resulting values of $\boldsymbol{v}_\odot \cdot \hat{\boldsymbol{d}}_{\rm MW}$ for each target galaxy. This quantity is the difference between its GRV and its HRV.

2.4 Comparing simulations with observations

Our simulations yield a velocity field for the LG. To determine the model-predicted GRV of an observed galaxy, we need a test particle landing at exactly the same position. To achieve this, we started with whichever test particle landed closest to the targeted final position. We then used a 2D Newton–Raphson algorithm on the initial position of this particle. The dependence of its final position on its initial one was found using finite differencing. For this, we used trajectories starting at (x_0, y_0), $(x_0 + dx_0, y_0)$ and $(x_0, y_0 + \mathrm{d}y_0)$, with $\mathrm{d}x_0 = \mathrm{d}y_0 = 307$ pc. Note that we have reverted to the usual xy coordinates, with $\hat{\boldsymbol{y}}$ pointing from M31 towards the MW and $\hat{\boldsymbol{x}}$ orthogonal to this direction.

We considered the Newton–Raphson algorithm to have converged once the error in the final position (x, y) was below 0.001 per cent of the distance between the target and the LG barycentre. The final velocity of this trajectory (v_x, v_y) was used to determine the model-predicted GRV of the target galaxy. We then corrected this for the motion of the Sun with respect to the MW to obtain its model-predicted HRV.

$$\mathrm{GRV}_{\rm model} = \frac{v_x x + (v_y - \dot{y}_{\rm MW})(y - y_{\rm MW})}{\sqrt{x^2 + (y - y_{\rm MW})^2}} \tag{50}$$

$$\mathrm{HRV}_{\rm model} = \mathrm{GRV}_{\rm model} - \boldsymbol{v}_\odot \cdot \hat{\boldsymbol{d}}_{\rm MW}. \tag{51}$$

If the MW or M31 mass is altered, then another simulation is required. However, if we only wish to alter the adopted $v_{c,\odot}$, then this is not necessary. We simply use the same $\mathrm{GRV}_{\rm model}$ but a different $\boldsymbol{v}_\odot$ (equation 47). In general, this alters $\mathrm{HRV}_{\rm model}$.

To account for distance uncertainties, the target was moved to the 1σ upper limit of its observed distance $d_{\rm MW}$ (using the 1σ lower limit instead had a negligible impact on our analysis). The Newton–Raphson procedure was then repeated targeting the revised position. Once this converged, we extracted the GRV from the final trajectory. We took the difference between these GRV estimates and called this $\sigma_{\rm pos}$. This is the uncertainty in the model-predicted GRV of a target galaxy due to uncertainty in its position.

$$\sigma_{\rm pos} \equiv \left|\mathrm{GRV}_{\rm model}\left(d_{\rm MW} + \sigma_{d_{\rm MW}}\right) - \mathrm{GRV}_{\rm model}\left(d_{\rm MW}\right)\right|. \tag{52}$$

Here, $\sigma_{d_{\rm MW}}$ is the uncertainty in the distance to a target galaxy. We assume negligible uncertainty in the direction towards it, constraining its position to be along a line. The velocity field is treated as linear over the part of this line where the target galaxy is likely be. Thus, assuming distance errors to be Gaussian, $\mathrm{GRV}_{\rm model}$ would also have a Gaussian distribution.

To determine $\sigma_{\rm pos}$ for M31, we use a slightly different procedure because it is not massless. Once we have the time history of the MW–Cen A separation, we keep this fixed and vary the initial MW–M31 separation to target a revised final value. For consistency, we also do not change $M(t)$. The effect on the final GRV of M31 is used for $\sigma_{\rm pos}$.

We expect this procedure to be approximately correct because Cen A only affects the GRV of M31 by $\sim 10\,\mathrm{km\,s^{-1}}$, making it not crucial to handle tides from Cen A very accurately. It would be possible to do so by recalculating trajectories for all three galaxies with revised target positions, but due to numerical difficulties this would probably have been less precise. It will become clear later that our results are not much affected by the value of $\sigma_{\rm pos}$ for M31.

Altering the MW–M31 separation changes the gravitational field in the RLG, affecting GRVs of objects within it. We expect this to be a very small effect and so we neglected it.

We conducted simulations across a wide range of total initial masses M_i and mass fractions in the MW q_1 (see Table 2). For situations with $q_1 > \frac{1}{2}$, we took advantage of a symmetry that arises between situations with $q_1 \leftrightarrow 1 - q_1$. Essentially, the behaviour of M31 in the low-q_1 case is equivalent to the behaviour of the MW in the high-q_1 case. Thus, we did not repeat all our calculations for the latter.

The positions of observed galaxies were altered in the following way:

$$x \rightarrow x \quad \text{(unaltered)} \tag{53}$$

$$y = y_{\rm M31} - y_{\rm rel} \ \text{ instead of } \left(y_{\rm MW} + y_{\rm rel}\right), \tag{54}$$

where $y_{\rm rel}$ is still obtained using equation (46) and is therefore unchanged. equation (54) also applies to Cen A, so its final position is now different. This meant we had to find a new solution for the trajectories of the MW, M31 and Cen A respecting the revised constraint on Cen A. Once this was done, we had to deal with altered positions (x, y) for our target galaxies by finding new test particle trajectories with the right final positions. The final GRVs of these trajectories were obtained using equation (50), but referred to M31 rather than to the MW.

The step we did not repeat was the calculation of the LG velocity field. This meant we had a much poorer guess for the initial position of each target galaxy. For this, we simply re-used the values of x and $y_{\rm rel}$ at the initial time in the low-q_1 case. Despite this, our algorithm still converged.

This procedure implicitly assumes that the accretion radii of the MW and M31 are swapped (i.e. that the galaxy with the higher mass always has the larger accretion radius). However, with the very low amounts of mass accreted by these galaxies (Fig. 6), this should hardly affect our results. This is especially true when considering that our analysis tends to disfavour $q_1 > \frac{1}{2}$ (Fig. 7).

As well as uncertainties due to position ($\sigma_{\rm pos}$) and measurement error on the radial velocity (σ_{v_h}), we also included an extra variance

term, $\sigma_{\rm extra}$. This was to account for effects not handled in our algorithm, for instance interactions between LG dwarf galaxies and tides raised by large-scale structures (LSS). $\sigma_{\rm extra}$ is a measure of how much model-predicted and actual radial velocities disagree. Including it, the contribution to χ^2 of any particular galaxy i is

$$\chi_i^2 \equiv \left(\frac{\mathrm{HRV_{model}} - \mathrm{HRV_{obs}}}{\sigma}\right)^2 \quad \text{where} \tag{55}$$

$$\sigma = \sqrt{\sigma_{\rm pos}^2 + \sigma_{v_h}^2 + \sigma_{\rm extra}^2}. \tag{56}$$

The uncertainty on the motion of the Sun can introduce systematic errors into our analysis. Thus, we treated $v_{c,\odot}$ as another model parameter. However, it is independently constrained ($239 \pm 5\,\mathrm{km\,s^{-1}}$; McMillan 2011). This was accounted for using a Gaussian prior, or equivalently by adding an extra contribution to χ^2. Therefore, the total χ^2 for any particular model (≡ combination of model parameters) is

$$\chi^2 = \sum_{\substack{\text{Target}\\ \text{galaxies}}} \chi_i^2 + \left(\frac{v_{c,\odot} - v_{c,\odot,\rm nominal}}{\sigma_{v_{c,\odot}}}\right)^2. \tag{57}$$

Models with higher $\sigma_{\rm extra}$ will necessarily achieve a lower χ^2. Thus, we cannot use χ^2 alone to decide which models are best. We made use of the fact that the probability of a model matching an individual observation

$$P(\text{Observation of galaxy } i \,|\, \text{Model}) \propto \frac{1}{\sigma} \mathrm{e}^{-\frac{\chi_i^2}{2}}. \tag{58}$$

Thus, we recorded both χ_i^2 and σ_i for each observed galaxy. The relative model likelihoods were then found using

$$P(\text{Model}|\,\text{Observations}) \propto \left(\prod_i \frac{1}{\sigma_i}\right) \mathrm{e}^{-\frac{\chi^2}{2}}. \tag{59}$$

If model-predicted and observed HRVs often disagree by much more than observational errors, then non-zero values of $\sigma_{\rm extra}$ will be preferred. Once χ^2 becomes comparable to the number of target galaxies (32), increasing $\sigma_{\rm extra}$ further will not much reduce χ^2. As a result, instead of P increasing with $\sigma_{\rm extra}$, it will actually start to *decrease* because of the factors of $\frac{1}{\sigma_i}$ in equation (59). One can imagine this as penalizing models where χ^2 is so small that such good agreement with observations is 'too good to be true'.

In this way, we hoped to constrain $\sigma_{\rm extra}$. If model-predicted and observed HRVs agree well given observational uncertainties, then the posterior distribution of $\sigma_{\rm extra}$ would peak at or near 0. If that does not occur, then this might indicate underestimated observational errors or a failure of the model.

Physically, we expect the main source of astrophysical noise contributing to $\sigma_{\rm extra}$ to be interactions between LG dwarf galaxies. However, Andromeda is much heavier than them, suggesting that it should be treated somewhat differently. This is because a minor merger would affect its velocity very little. Thus, whatever the adopted value of $\sigma_{\rm extra}$ for other LG galaxies, a smaller value of $\sigma_{\rm extra,\,M31}$ should be adopted for M31. We used $\frac{\sigma_{\rm extra,\ M31}}{\sigma_{\rm extra}} = 0.1$. This alters equation (56) for M31 and thus its contribution to χ^2.

We considered the effect of a minor merger with Andromeda or the MW in the past. This was modelled as an impulse, meaning that we instantaneously altered the GRV of M31 at some time in the past. The effect on its present GRV was then determined. For simplicity, Centaurus A was omitted and the total LG mass was held constant at $4 \times 10^{12}\,\mathrm{M_\odot}$. This roughly reproduces the present GRV and distance of M31.

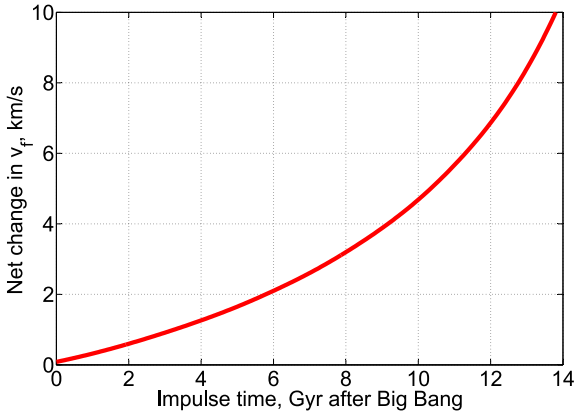


Figure 4. The overall effect on the present GRV of Andromeda due to a $10\,\mathrm{km\,s^{-1}}$ impulse to its GRV in the past, with the present distance to Andromeda constrained. This constraint is maintained by altering the initial MW−M31 separation (see text). Because of this, impulses applied longer ago have a smaller net effect on present motions.

One might think that the longer ago the impulse was, the bigger its effect on the present GRV of M31, Δv_f. After all, pushing the galaxies towards each other increases the force between them at later times, further reinforcing the original impulse.

However, this would lead to the constraint on the present distance to M31 being violated (in this example, it would end up too close). Consequently, we had to alter the initial separation of the galaxies d_i compared with a non-impulsed trajectory. This tends to counteract the direct effect of the impulse.

The results we obtained for Δv_f as a function of the impulse time are shown in Fig. 4. An impulse applied very recently hardly affects d_f and so d_i does not have to be altered much. Thus, Δv_f is almost equal to the impulse.

For impulses applied longer ago, Δv_f rapidly becomes very small. The dependence on impulse time is even steeper than for Hubble drag ($\Delta v_f \propto a^{2.4}$, where a was the cosmic scalefactor when the impulse was applied). This underlines just how difficult it is to alter the present GRV of M31. Consequently, a realistic model needs to match this constraint very well.

As well as interactions between LG dwarf galaxies, our model does not fully account for the presence of LSS in the Universe beyond the LG. We attempted to include some of these structures in Section 4.3, but others remain beyond the scope of this investigation. The leading order effect of LSS on the LG is to accelerate it as a whole without altering the relative velocities between objects within it. However, LSS also raise tides on the LG, affecting the GRVs of our target galaxies. Such effects are larger for galaxies further from the MW. As M31 is the closest galaxy in our sample, its GRV should be least affected by tides raised by LSS. This further justifies our decision to use a value for $\frac{\sigma_{\rm extra,\,M31}}{\sigma_{\rm extra}}$ that is much smaller than 1.

3 RESULTS

Our analysis works by determining the model-predicted GRV of each target galaxy in the LG. A range of models are tried out, with different initial LG masses M_i and mass fractions in the MW q_1 (see Table 2). The results for two target galaxies are shown in Fig. 5. Each of these GRV predictions are converted into a range of HRV predictions using $v_{c,\odot}$ within 3σ of its most likely value

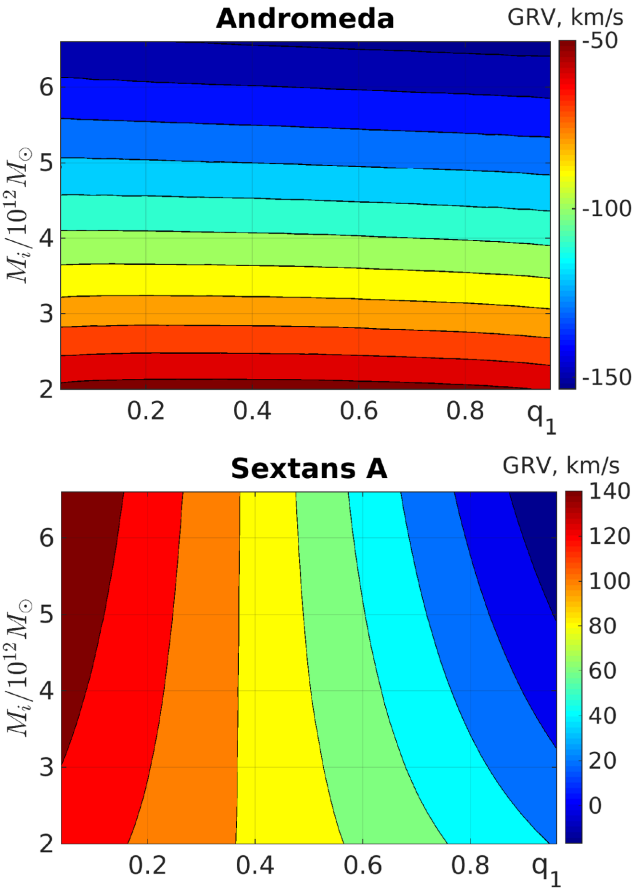


Figure 5. Simulated GRVs as a function of model parameters for two target galaxies. Different galaxies constrain a different combination of model parameters, with Andromeda mostly telling us the total LG mass and Sextans A mostly telling us the fraction of this mass in the MW (because the GRV of Sextans A is much higher at low q_1).

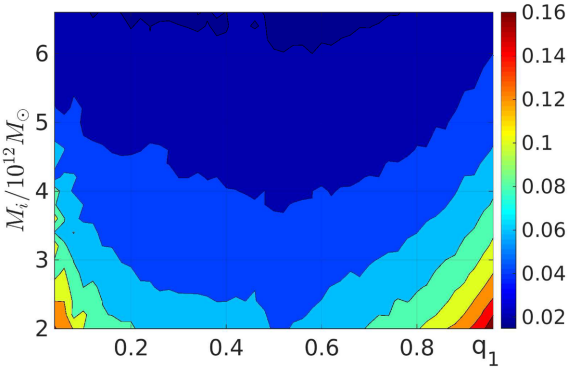


Figure 6. Fraction of the initial MW mass that it accretes by the end of our simulations. Due to symmetry, approximate results for Andromeda can be obtained by setting $q_1 \to 1 - q_1$.

according to the work of McMillan (2011). By comparing these with observed HRVs, we obtain complementary constraints on the model parameters. As we have > 2 target galaxies, we also test the model itself.

Our simulations allowed the MW and M31 to accrete mass. In Fig. 6, we show the fraction by which the original mass of each galaxy increased. The galaxies only increase their mass by a few per cent in our simulations. Thus, accretion is unimportant in them. This is mainly because a test particle needs to pass within a few disc scalelengths of the MW or M31 for us to consider the particle accreted (Table 2).

This aspect of our models is not totally realistic. If more distant approaches were also treated as leading to accretion, then the MW and M31 would gain more mass. We do not consider this an important effect because we tried a wide range of initial masses for both galaxies.

Fig. 7 shows the posterior probability distributions of all our model parameters and pairs of parameters based on a set of 1128 simulations[14] that include Centaurus A with a mass of $4 \times 10^{12}\,\mathrm{M}_\odot$. Each simulation was compared with observations using 101 values of $v_{c,\odot}$ and 201 values of $\sigma_{\rm extra}$ (priors are given in Table 2). Of particular importance is the posterior on $\sigma_{\rm extra}$, which we constrain to be $45^{+7}_{-6}\,\mathrm{km\,s^{-1}}$. As observational errors are typically ∼5–10 km s^{-1} and are already included in our analysis, this is very surprising.

We checked if varying the start time of our simulations from $a_i = \frac{1}{10} \to \frac{1}{15}$ affected our results. This reduced the most likely value of $\sigma_{\rm extra}$ by $\sim 1\,\mathrm{km\,s^{-1}}$. Apparently, our results are not much affected by the epoch at which our simulations are started. Some reasons for this are given in Section 4.6).

We considered a different estimate for the LSR speed ($238 \pm 9\,\mathrm{km\,s^{-1}}$; Schönrich 2012). As might be expected, this affected $\sigma_{\rm extra}$ by $\lesssim 1\,\mathrm{km\,s^{-1}}$. This is because we consider $v_{c,\odot}$ to be well constrained independently of our work. It is also apparent that there is very little tension between these independent measurements and our timing argument analysis (Table 2).

Our special treatment of M31 forces up $\sigma_{\rm extra}$ to some extent as it essentially forces our models to match its GRV (given the small uncertainty on $v_{c,\odot}$). As this may be overly restrictive, we redid our analysis using the same value of $\sigma_{\rm extra}$ for M31 as for other LG galaxies (i.e. $\frac{\sigma_{\rm extra,M31}}{\sigma_{\rm extra}} = 1$ instead of 0.1). This lowered $\sigma_{\rm extra}$ by $\sim 2\,\mathrm{km\,s^{-1}}$.

Our method of handling distance uncertainties is very similar to that used by Peñarrubia et al. (2014). We rely on an assumption that the velocity field is approximately linear over the range of positions where the galaxy could plausibly be. To test this, we repeated our calculations with $\sigma_{\rm pos}$ estimated based on how much the simulated GRV of each target galaxy changed if we altered its distance from its observed value to its 1σ lower limit. This gave almost identical results to when we used the 1σ upper limit instead ($\sigma_{\rm extra}$ decreased by $\sim 0.1\,\mathrm{km\,s^{-1}}$ when using the lower limit).

Our analysis favours a very low value for q_1, the fraction of the LG mass originally in the MW. This is related to the fact that observed HRVs tend to systematically exceed the predictions of the best-fitting model (Fig. 9). Thus, our analysis will prefer those models that generally lead to increased GRVs. Reducing q_1 has this effect because it causes particles projected orthogonally to the MW–M31 line; to curve towards M31 and away from the MW. It also implies a faster motion of the MW relative to the LG barycentre and a greater distance from there, enhancing projection effects.

Most of our target galaxies are in fact roughly orthogonal to the MW–M31 line as perceived from the LG barycentre (Fig. 3). This might be why q_1 seems to have a strong impact on GRVs (bottom panel of Fig. 5). Thus, one might expect our analysis to prefer very low values of q_1, which indeed it does.

Certain correlations are apparent between some of our model parameters. Because we require our models to accurately match the

[14] Spanning a linear grid with 24 steps in M_i and 47 steps in q_1, although some shortcuts were taken for $q_1 > \frac{1}{2}$.

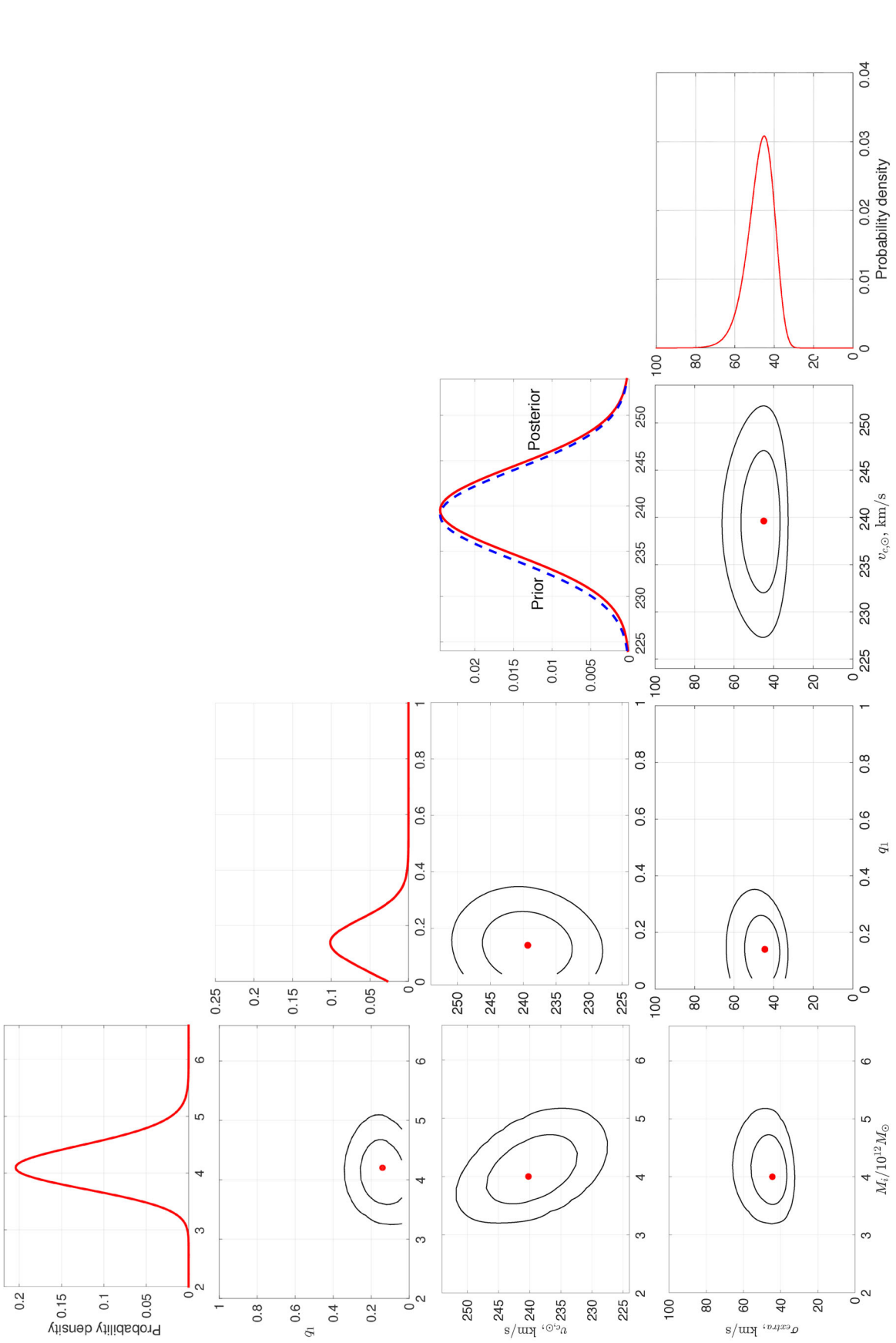


Figure 7. Posterior probability distributions of the parameters considered in this work (defined in Table 2). In each figure, we marginalized over the parameters not shown. For variables plotted against other variables, we show contours of the probability density corresponding to the 1σ (68.3 per cent) and 2σ (95.4 per cent) confidence levels, as well as the most likely pair of values. Rotate 90° clockwise for viewing.

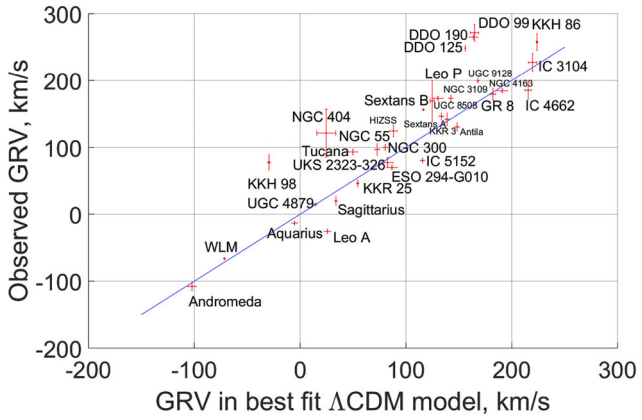


Figure 8. Comparison between predicted and observed GRVs based on the most likely model parameters ($q_1 = 0.14$, $M_i = 4.2 \times 10^{12}\,\mathrm{M}_\odot$, $v_{c,\odot} = 239\,\mathrm{km\,s^{-1}}$). The line of equality is in blue.

observed HRV of M31, our timing argument estimate of the LG mass is quite sensitive to anything which affects its predicted HRV. M31 is almost directly 'ahead' of the Sun in its motion around the MW. Thus, for the same GRV of M31, increasing $v_{c,\odot}$ decreases its HRV (equation 49). To increase its HRV back up to its observed value, its GRV would have to be increased, which is only possible in a different model where the retarding effect of gravity is smaller (i.e. M_i is lower).

A lot of our target galaxies have HRVs which exceed the predictions of the best-fitting model (Fig. 8). This means that a lower LG mass fits the data slightly better, explaining the correlation between M_i and σ_{extra}. For the same reason, increasing $v_{c,\odot}$ indirectly improves the fit to the data, reducing σ_{extra} slightly.

Some effects are inevitably not considered in our model. If they were included, we might achieve a better fit to the observations. We consider some of these effects in the next section. We pay special attention to tides from objects beyond the LG (Section 4.3) and the LMC (Section 4.4).

4 DISCUSSION

Our analysis reveals an astrophysical noise in velocities of LG galaxies that greatly exceeds observational errors. With $N = 32$ targets, the fractional uncertainty in this extra noise σ_{extra} should be $\approx \frac{1}{\sqrt{2N}} = \frac{1}{8}$. This agrees closely with the width of the posterior probability distribution of σ_{extra} (Fig. 7).

We considered several factors which could influence our analysis. Perhaps most obviously, the LG contains gravity from objects other than the MW and M31. For example, the non-satellite LG galaxies that we modelled as test particles in reality exert gravity on each other. This would lead to roughly isotropic and random impulses on them. Considering that our analysis is based solely on line of sight velocities, we would need to assume typical impulses of $\sim\sqrt{3}\,\sigma_{\mathrm{extra}} \approx 80\ \mathrm{km\ s^{-1}}$.

However, our target galaxies have typical velocity dispersions/rotation speeds of $\lesssim 15\ \mathrm{km\ s^{-1}}$ (e.g. Kirby et al. 2014). For some impact parameters, these galaxies could perhaps impulse each other by twice this while avoiding a merger. Thus, the high value of σ_{extra} inferred by our analysis seems difficult to explain as a result of interactions amongst the galaxies we considered.

Additional inaccuracies in our model may arise from the effects of LSS. Moreover, even distant encounters between LG dwarf galaxies can affect their motion. The likely magnitude of such effects can be estimated based on more detailed cosmological simulations of the ΛCDM paradigm. Considering analogues of the LG in such simulations, it has been found that the dispersion in radial velocity with respect to the LG barycentre at fixed distance from there should be $\sigma_{\mathrm{H}} \sim 30\ \mathrm{km\ s^{-1}}$ (Aragon-Calvo, Silk & Szalay 2011).

Table 3. Contributions of various sources to σ_{H}, the radial velocity dispersion with respect to the LG barycentre at fixed distance from there. The GA leads to an $\sim 40\,\mathrm{km\,s^{-1}}$ range in radial velocities at 3 Mpc (equation 62). Combining everything in quadrature, we can account for $\sim 20\,\mathrm{km\,s^{-1}}$ of the $\sim 30\,\mathrm{km\,s^{-1}}$ dispersion in the Hubble flow found by Aragon-Calvo et al. (2011). This suggests that our model should represent ΛCDM to an accuracy of $\sim 20\,\mathrm{km\,s^{-1}}$.

Object	Contribution to σ_{H} ($\mathrm{km\ s^{-1}}$)	Comments	Section
MW, M31 and Cen A	~ 15	10 at 2 Mpc	2.2
IC 342 and M81	~ 5	Table 6	4.3.2
The GA	~ 10	equation (62)	4.3.3

Looking at the bottom panel of Fig. 3, it is clear that our models do not produce such a large velocity dispersion. We seem to get $\sigma_{\mathrm{H}} \sim 10\ \mathrm{km\ s^{-1}}$, though this rises slightly to $\sim 15\,\mathrm{km\,s^{-1}}$ once we include Centaurus A. Thus, even if ΛCDM were correct, it would be reasonable for our analysis to infer $\sigma_{\mathrm{extra}} \sim 25\,\mathrm{km\,s^{-1}}$.

In this section, we hope to correct some of our model deficiencies and make it a more accurate representation of ΛCDM. Table 3 shows some of the effects we consider and a rough idea of their contributions to σ_{H}. Combining everything in quadrature suggests that the objects we consider are sufficient to attain a dispersion in the Hubble flow of $\sim 20\ \mathrm{km\ s^{-1}}$. Thus, a lot of the 'scatter' about the Hubble flow found by Aragon-Calvo et al. (2011) arises because the LG is not spherically symmetric rather than actually being a dispersion in velocities at the same position. None the less, another $\sim 20\,\mathrm{km\,s^{-1}}$ must come from factors we do not consider. This means that values of σ_{extra} much greater than $\sim 20\,\mathrm{km\,s^{-1}}$ would be problematic for ΛCDM.

Our best-fitting model has $q_1 = 0.14$. Results from this model are compared with observations in Fig. 8. However, we consider it unrealistic for the MW to have only $\frac{1}{6}$ as much mass as M31. Assuming that the virial mass of a halo scales as the cube of its velocity dispersion (Evrard et al. 2008) and that the ratio of the latter between the MW and M31 is $\frac{225}{180} = 1.25$ (Carignan et al. 2006; Kafle et al. 2012), we see that it is unlikely for M31 to have much more than twice as much mass as the MW. We believe the best compromise between this argument and the low value of q_1 preferred by our timing argument analysis (0.14 ± 0.07) is found if we set $q_1 = 0.2$. Thus, when comparing our model predictions with observations (Fig. 9 onwards), we use the model parameters which best fit the data but with q_1 raised to 0.2. This raises σ_{extra} by $\sim 1\ \mathrm{km\ s^{-1}}$ and has only a small impact on our results, but should make them more realistic.

Observed GRVs seem to systematically exceed model predictions (Fig. 8). We used our most plausible model including Cen A to subtract model-predicted radial velocities from observed ones, yielding ΔGRV for each target galaxy. We then created a histogram of the resulting ΔGRVs in Fig. 9, smoothing each data point over its respective uncertainty. As before, this includes σ_{pos} and σ_{v_h}. Because it is unclear exactly how to convert heliocentric radial velocities into Galactocentric ones, we also added $\sigma_{v_{c,\odot}}$ in quadrature to all the uncertainties.

If one assumes that factors outside our model are just as likely to raise GRVs of target galaxies as to reduce them, then it should

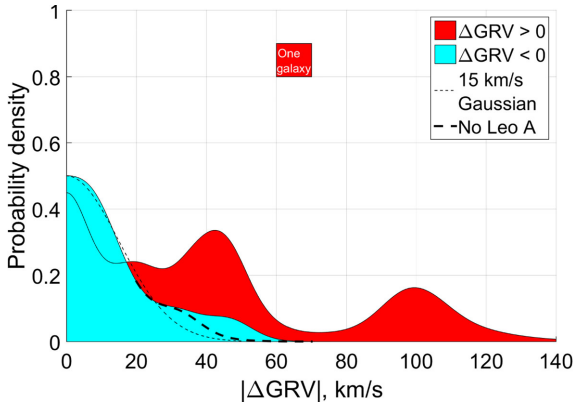


Figure 9. Histogram showing observed – predicted GRVs (i.e. ΔGRVs) of our target galaxies using our most plausible model ($q_1 = 0.2$ instead of 0.14, other parameters as in Fig. 8). Each data point was convolved with a Gaussian of width $\sigma = \sqrt{{\sigma_{\rm pos}}^2 + {\sigma_{v_h}}^2 + {\sigma_{v_{c,\odot}}}^2}$. We divided our sample into those with ΔGRV < 0 (blue) and those with ΔGRV > 0 (red). The area corresponding to one galaxy is shown as a red square. A Gaussian of width 15 km s^{-1} is overplotted as a short-dashed line. This matches the ΔGRV < 0 subsample quite well, especially when Leo A is excluded (long-dashed line).

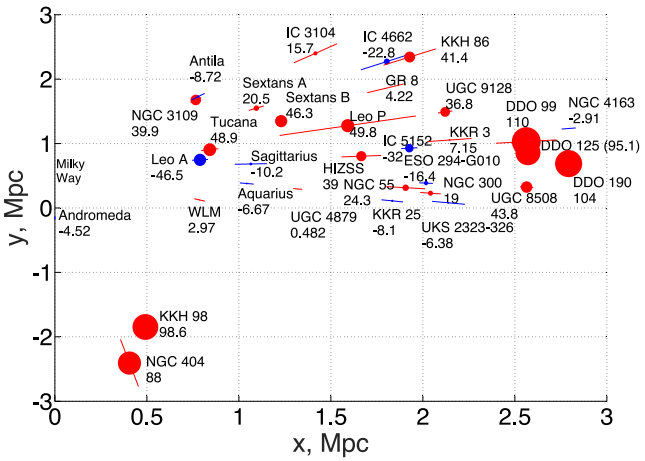


Figure 10. Positions of target galaxies are shown in the same way as the top panel of Fig. 3. Distance uncertainties are indicated by a thin line. The size of each marker is directly proportional to |ΔGRV| for the associated galaxy (name beside marker) based on the same model as shown in Fig. 9, with colour indicating sign (red means positive). Uncertainties in ΔGRV are roughly proportional to that in distance. For Leo P, this causes a 30 km s^{-1} uncertainty, though typical values are much smaller. We expect our model to represent ΛCDM to an accuracy of ∼ 25 km s^{-1} (see text), roughly the same as ΔGRV of NGC 55.

be possible to use the population of ΔGRV < 0 galaxies to gain a good idea of how accurately our model represents ΛCDM. The ΔGRV < 0 galaxies are well described by a Gaussian of width 15 km s^{-1} (blue area in Fig. 9). Most of the mismatch is due to Leo A. To account for tides raised by IC 342 and M81, its radial velocity prediction should be reduced by ∼ 5 km s^{-1}, making it more consistent with observations (Table 6). In any case, considering the ΔGRV < 0 galaxies suggests that inaccuracies in our model probably do not exceed 25 km s^{-1}, slightly less than the $\sigma_{\rm H}$ ∼ 30 km s^{-1} found by Aragon-Calvo et al. (2011) due to our careful modelling. Thus, one might expect a Gaussian of around this width to also describe the distribution of ΔGRVs for galaxies with ΔGRV > 0.

However, unlike galaxies with ΔGRV < 0, those with ΔGRV > 0 are not well described by a 15 km s^{-1} Gaussian (red area in Fig. 9). There appears to be a population of ΔGRV > 0 galaxies which might be described by such a Gaussian, but in this case we would need perhaps two additional populations to fully account for the observations. A possible mechanism for generating these populations is described in Section 4.6.

We tried to see if there was any correlation between the position of a target galaxy and its associated ΔGRV. This is shown in Fig. 10, with the size of the marker for each galaxy directly proportional to its ΔGRV. It is assumed that the LG is axisymmetric, so positions are shown using the same coordinate system as our simulations. The uncertain distance to each galaxy is indicated by a thin line.

The objects with the highest ΔGRVs tend to be furthest from the MW/M31. This might be a sign that tides from objects outside the LG are responsible for the discrepancies. As we already included Centaurus A in our simulations, we might be seeing the effects of other objects. We will investigate some possibilities in Section 4.3. In particular, we will show that IC 342 and M81 are unlikely to be responsible for the discrepancies (Section 4.3.2). This is also true of the GA (Section 4.3.3). An explanation for this trend is suggested in Section 4.6.

4.1 Reduced LG mass

Comparison with cosmological simulations suggests that the timing argument may overestimate the LG mass (González et al. 2014). Moreover, observed radial velocities tend to be systematically more outwards than in our models. These considerations suggest that a lower LG mass could help to explain the observations. To test this, we removed the effect of gravity altogether and used equation (1) to predict velocities. As before, we took the barycentre of the LG as the centre of expansion and assumed that $v_{c,\odot} = 239$ km s^{-1}.

The MW was assumed to be going towards this point at ∼90 km s^{-1}, which is reasonable given the observed HRV of M31 and a plausible mass ratio between the galaxies. In theory, the MW should be going away from the LG barycentre in the absence of gravity. However, using the correct MW velocity ensures that velocities with respect to the LG barycentre are correctly converted into velocities with respect to us. One slightly unusual consequence of this is that the model predicts M31 to have a negative GRV.

GRV predictions obtained in this way are compared with observations in Fig. 11. Due to the effect of gravity, observed GRVs tend to be less than in a pure Hubble flow. Surprisingly, this is not true for some of our target galaxies, especially the DDO objects. Other examples of this behaviour have been identified recently (Pawlowski & McGaugh 2014a). Thus, reducing M_i does not explain the observations, at least if considered on its own.

Moreover, there is limited scope to alter M_i because it is tightly correlated with the present GRV of M31 (Fig. 5). A model needs to match its GRV fairly well because it is unlikely that a minor merger with M31 or the MW could have substantially affected their relative motion. Even if such an event did occur, its net effect on the present GRV of M31 would be greatly diminished unless it occurred recently (Fig. 4).

In our models, the mass in the LG is present not only in the MW and M31. We assume that the RLG contains a uniform distribution of matter with the same density as the mean cosmic density of matter $\overline{\rho}$. At this density, a sphere of comoving radius 2.9 Mpc would have a mass equal to that of the MW and M31, assuming $M_{\rm MW+M31} = 4 \times 10^{12}\,{\rm M}_\odot$.

At early times, it is possible that a small region encompassed the material that would later end up in the MW and M31. A surrounding

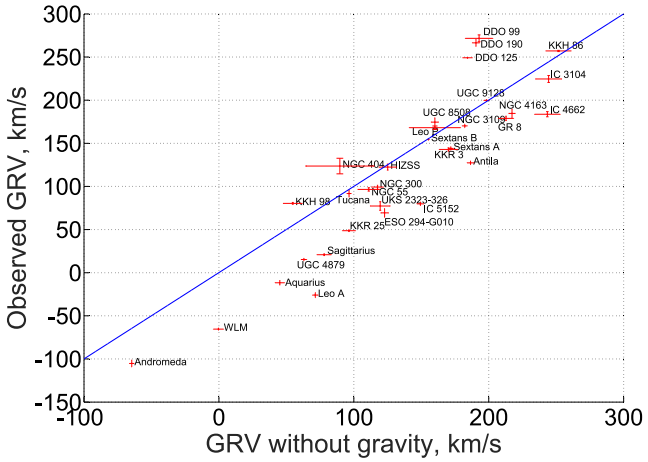


Figure 11. Comparison between observed GRVs and the predictions of equation (1), with the centre of expansion at the barycentre of the LG assuming $q_1 = 0.2$. The MW is taken to be moving towards this point at 90 km s^{-1}. Note that some galaxies are moving outwards even faster than a pure Hubble flow (blue line).

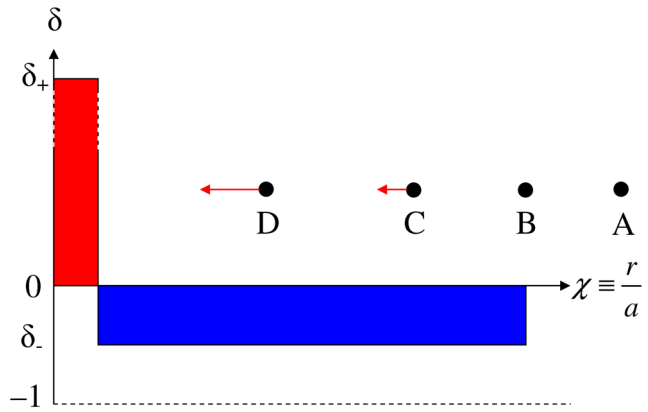


Figure 12. This shows a simple 1D model for the fractional overdensity δ at the start of our simulations, plotted in comoving coordinates χ. The MW and M31 are treated as one object which formed from the over-density at low χ. The overdensity must be surrounded by an underdense region so that the two cancel out on large scales (i.e. $\int \delta(\chi)\chi^2 d\chi = 0$). To avoid negative density, $\delta_- > -1$. We also show some test particle trajectories based on solving equation (60). Particles starting further out than B simply trace the cosmic expansion (χ is constant) because the mass enclosed interior to their radius is the same as in a homogeneous universe. This is not true for particle C – the region it encloses is overdense, so χ of particle C decreases slightly. The effect is larger for particles starting closer to the LG barycentre (e.g. D).

underdense region would be required so that the mean density in the union of both regions was $\overline{\rho}$. This is depicted schematically in Fig. 12.

If we assume the underdense region was completely empty (i.e. $\delta_- = -1$), then it would have to extend out to a comoving radius of 2.9 Mpc. As a result, there would be no mass in the RLG, assuming this was defined to have a radius below 2.9 Mpc. It can be seen from Fig. 3 that all our target galaxies have distances from the LG barycentre of <2.9 Mpc. Thus, they could all be in a void.

However, one must bear in mind that test particles are retarded by the gravity of the MW and M31. This reduces the comoving volume spanned by a cloud of test particles. Thus, if the RLG is not completely empty, then the LG contains material initially outside its present comoving volume.

To investigate the interplay between these effects, we now consider the opposite limit in which $|\delta_-| \ll 1$. This corresponds to a much larger underdense region surrounding the MW and M31 at the start of the simulations (∼500 million years after the big bang). To better understand this case, we solved some test particle trajectories assuming a point mass in an otherwise homogeneous universe. We kept fixed the mass enclosed interior to the radius of any given test particle. This makes the equation of motion[15]

$$\ddot{r} = -\frac{GM_{\rm eff}}{r^2} + {H_0}^2 \Omega_{\Lambda,0} r \tag{60}$$

$$M_{\rm eff} = M + \frac{4\pi}{3} {r_i}^3 \overline{\rho}_i \, (1 + \delta_-) \text{ (note : } \delta_- < 0), \tag{61}$$

where $\overline{\rho}_i$ is the mean density of matter in the Universe at the time our simulations are started. $M_{\rm eff}$ includes both the point mass M and any material originally present at radii below the initial radius of the test particle. We assume that $M_{\rm eff}$ *remains constant* because the system avoids crossing of shells. This can be achieved if the massive object accretes any objects that come sufficiently close to it, rather than just letting them escape on the other side.

To obtain a final distance from the LG of 2.9 Mpc, we need an initial distance of 0.46 Mpc for a starting time corresponding to when $a_i = 0.1$. This means that 16×10^{12} M$_\odot$ would end up within the RLG at the present time. If the RLG were to contain matter at $\overline{\rho}$, then it would only contain 4×10^{12} M$_\odot$. Thus, the RLG might have up to $\left(\frac{0.46}{0.29}\right)^3 = 4$ times as much material as was assumed in our calculations.

It is difficult to know how much mass is actually present in the RLG. There might be a diffuse component of dark matter or concentrations of it that have no detectable stars. Some regions are difficult to survey because of e.g. the disc of our Galaxy. Recently, significant amounts of hot gas have been discovered around the MW (Salem et al. 2015) and around M31 (Lehner, Howk & Wakker 2015).

For these reasons, we assumed neither of the extreme cases just outlined. Instead, we used an intermediate assumption that the RLG contains matter with a mean density of $\overline{\rho}$ and little density variation. Roughly speaking, this corresponds to $\delta_- = -0.43$ and an underdensity out to 3.5 Mpc. We think this is reasonable considering the distances to major mass concentrations just outside the LG (Table 1).

We investigated whether altering this assumption might affect our conclusions regarding the inferred value of $\sigma_{\rm extra}$. To do this, we assumed the extreme case that the RLG has no mass. This means that the $\frac{\ddot{a}}{a}$ term present in the equations of motion (e.g. equation 32) should be replaced with ${H_0}^2 \Omega_{\Lambda,0}$. We re-ran our entire analysis using equations of motion altered in this way.

If we used the same procedure as before to prevent test particles starting too close to the MW/M31, then we would end up with no test particles within ∼2.9 Mpc of the LG barycentre. In this case, there would be no way to obtain HRV predictions for our target galaxies, consistent with the assumption of an empty RLG. Clearly, this assumption is wrong at some level. Thus, we allow test particle trajectories to start anywhere as long as they end up at the correct position.

[15] It can be verified that a pure Hubble flow is recovered for the case $M = \delta_- = 0$.

Table 4. Alterations to cosmological parameters in Table 2 for the high H_0 model shown in Fig. 13. In both cases, the universe is flat and equally old (13.81 Gyr).

Parameter	Old value	New value
H_0	67.3 km s^{-1} Mpc^{-1}	72.3 km s^{-1} Mpc^{-1}
$\Omega_{\rm m,\,0}$	0.315	0.243
$\Omega_{\Lambda,\,0}$	0.685	0.757

Altering the setup of our simulations in this way reduced $\sigma_{\rm extra}$ by $\sim$7 km s^{-1}. Individual radial velocities are often increased by larger amounts. This tends to reduce the discrepancy with observations. However, the overall effect is small because, for the same MW and M31 mass, the GRV of M31 is increased. To bring it back down to the observed value, the mass of the MW and M31 have to be increased, reducing the predicted GRVs of other LG galaxies.

We checked this explicitly by comparing the marginalised posterior probability distribution of M_i. Assuming the RLG has a mean density of $\overline{\rho}$, the total initial LG mass in units of $10^{12}\,{\rm M}_\odot$ is 4.2 ± 0.4. However, if we assume an empty RLG, this rises to 5.2 ± 0.4. This seems rather high, but is in line with similar calculations by other workers (Partridge et al. 2013).[16] We suggest that this result points towards a RLG that cannot be considered empty for the purposes of the timing argument. However, more reasonable values for M_i are obtained if one includes the kinematic effect of a sufficiently massive LMC (Fig. 18).

4.2 Increased Hubble constant

Another way to increase model-predicted HRVs is to increase H_0. The cosmological value seems to be fairly well constrained (Planck Collaboration XIII 2015). Once certain biases are taken into account, this measurement seems to be consistent with surveys of Type Ia supernovae (Rigault et al. 2015). However, there is also some cosmic variance: underdense regions of the Universe expand faster than the average. If we are in such a region, this would lead to the value of H_0 appropriate for the local Universe being higher than that for the Universe as a whole (Wojtak et al. 2014).

To account for this possibility, we performed another simulation with H_0 raised by 5 km s^{-1} Mpc^{-1}. However, we were careful to bear in mind that Planck gives a tight constraint on the age of the Universe. To avoid altering this, we had to further adjust the adopted cosmology. For simplicity, we kept this flat. The parameters used are shown in Table 4.

The resulting posterior on $\sigma_{\rm extra}$ is shown in Fig. 13. As might be expected, increasing H_0 lowers the inferred value of $\sigma_{\rm extra}$, but only by $\sim$5 km s^{-1}. This is similar to the effect of assuming the RLG is empty instead of filled with matter at a density of $\overline{\rho}$ (Section 4.1). This is reassuring as the simulations work in slightly different ways.

Of course, it is only possible to count this reduction in $\sigma_{\rm extra}$ once: to account for the RLG being less dense than in our models, one can either alter the equations of motion to make the RLG empty or one can raise H_0 slightly. Whichever method one prefers to use, the effect is not sufficient to explain the observations, although it does help.

One thing that may be in favour of models with an underdense RLG is the inferred value of q_1, the fraction of the LG mass in the MW. When we tried to make the RLG empty by altering the equations of motion (Section 4.1), our analysis preferred $0.38^{+0.06}_{-0.05}$ instead of 0.14 ± 0.07. One might expect a similar effect to occur when we raise H_0 – after all, the effect on $\sigma_{\rm extra}$ is very similar. However, our calculations show almost no change in the inferred value of q_1 due to a higher Hubble constant.

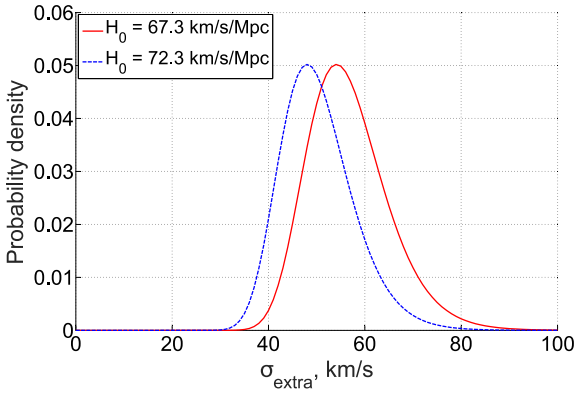


Figure 13. The posterior on $\sigma_{\rm extra}$ is shown for two plausible values of the Hubble constant H_0. The age of the universe is the same in both models, with parameters adjusted accordingly (Table 4). Raising H_0 by 5 km s^{-1} Mpc^{-1} reduces $\sigma_{\rm extra}$ by $\sim$5 km s^{-1}.

Table 5. Effect of Centaurus A on posteriors of model parameters. Both analyses shown here have a uniform prior on q_1 over the range 0.2–0.8.

Parameter and units	Prior	Posterior without Cen A	Posterior with Centaurus A
$\sigma_{\rm extra}$, km s^{-1}	0–100	$54^{+8.9}_{-7.0}$	$46^{+7.4}_{-5.6}$
M_i, $10^{12}\,{\rm M}_\odot$	2–6.6	$3.42^{+0.35}_{-0.32}$	$4.14^{+0.35}_{-0.31}$
q_1	0.2–0.8	$0.20^{+0.060}_{-0}$	$0.20^{+0.049}_{-0}$
$v_{c,\,\odot}$, km s^{-1}	239 ± 5	$242.0^{+4.9}_{-4.7}$	$239.9^{+4.9}_{-4.7}$

4.3 Tides from objects outside the LG

4.3.1 Centaurus A

To better understand the effect of Cen A on our results, we repeated our analysis without including it. The results are shown in Table 5. Broadly speaking, the results are similar in both cases, although there are some subtle differences.

Being close to the MW–M31 line, Cen A pulls the MW and M31 apart, increasing the GRV of M31 by $\sim$10 km s^{-1}. To bring it back down to the observed value, M_i would need to be increased by $\sim 10^{12}\,{\rm M}_\odot$ (Fig. 5). This is indeed roughly what happens to the posterior on M_i. Other effects are harder to understand, such as why including Cen A leads to better agreement with the LSR speed measured by McMillan (2011).

The posterior distribution of $\sigma_{\rm extra}$ is shown in Fig. 14. Including Cen A reduces its most likely value from 54^{+9}_{-7} km s^{-1} to 46^{+7}_{-6} km s^{-1}. If H_0 is also increased to 72.3 km s^{-1} Mpc^{-1}, then $\sigma_{\rm extra}$ is further reduced to 41^{+6}_{-5} km s^{-1}.

We only tried one possible mass of Cen A ($4 \times 10^{12}\,{\rm M}_\odot$). Including it at this mass reduces $\sigma_{\rm extra}$ by $\sim$ 8 km s^{-1}. It is possible that adopting a higher mass would reduce it further.[17] However, it is unlikely that Cen A is more massive than $5 \times 10^{12}\,{\rm M}_\odot$ (see

[16] Part of the difference arises because Cen A is not usually included in timing argument analyses of the LG.

[17] Though it might not, see Fig. 17.

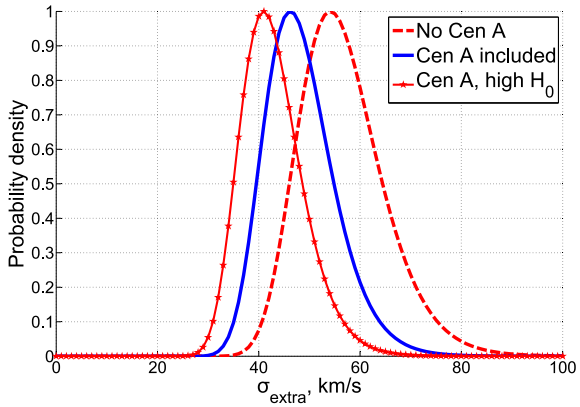


Figure 14. Posterior probability distribution of $\sigma_{\rm extra}$ under different model assumptions. Including Centaurus A reduces $\sigma_{\rm extra}$ by $\sim$8 km s^{-1}. Also raising H_0 as described in Table 4 reduces $\sigma_{\rm extra}$ by a further $\sim$5 km s^{-1}.

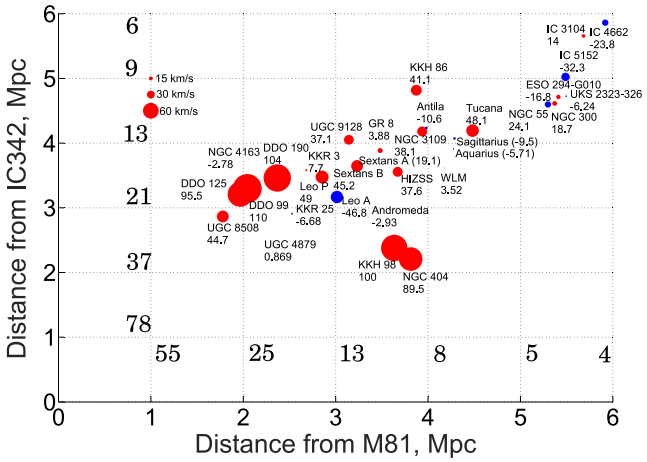


Figure 15. Observed – predicted HRVs (i.e. ΔHRVs) of indicated galaxies as a function of distance from the perturbers in Table 1. The radius of each marker ∝ the discrepancy, which we list below the name of each galaxy. Blue indicates a measured HRV below that in the most plausible model ($q_1 = 0.2$, $M_i = 4.2 \times 10^{12}$ M$_\odot$), while red shows the opposite. The numbers on the gridlines show the peculiar velocity in km s^{-1} at that distance from each perturber if it was the only object in the universe (see text). A more detailed model for how perturbers affect observations is shown in Table 6.

fig. 1 of Karachentsev 2005). This is only 25 per cent higher than our adopted mass. Thus, using the highest plausible Cen A mass rather than our adopted value might well reduce $\sigma_{\rm extra}$, but only by another $\sim$ 2 km s^{-1}.

The inclusion of Centaurus A affects the Hubble diagram for the LG, increasing $\sigma_{\rm H}$. This is a tidal effect and is therefore larger at greater distances. At 3 Mpc, we found a range in radial velocity from the LG barycentre of $\sim$ 70 km s^{-1}, falling to perhaps half that at 2 Mpc. This corresponds to $\sigma_{\rm H} \sim 10-15$ km s^{-1}. Although this is below the $\sim$30 km s^{-1} found in cosmological simulations (Aragon-Calvo et al. 2011), it does suggest that some of the 'scatter' about the Hubble flow found in such simulations can be accounted for using an axisymmetric model rather than a spherically symmetric one.

4.3.2 *IC 342 and M81*

Including Centaurus A improves the fit to the data slightly but still leaves a very poor fit. As it is the most massive perturber, this suggests that tides cannot explain the discrepant HRVs. To check this conclusion, we cross-correlated the discrepancies in the HRVs with the distances between our target galaxies and the remaining perturbers in Table 1. This is shown in Fig. 15. The discrepancy seems to be larger for objects closer to IC 342 or to M81. Thus, we tried to see if tides from these objects might help to explain the observations.

We provide two ways of estimating the effects of tides raised by IC 342 and M81 on the LG. First, we treat each perturber as the only object in the universe. We solve test particle trajectories in the usual way and target a particular final separation with the perturber. We then record the peculiar velocity of this trajectory. Using the perturber masses in Table 1, the results obtained in this way are indicated in km s^{-1} on the gridlines of Fig. 15.

In this very simplistic model, the effect of each perturber is just an extra velocity towards it with the calculated magnitude. However, the direction of this velocity is not directly away from the MW. For example, IC 342 should hardly affect the GRV of KKH 98 because, as perceived by KKH 98, the MW and IC 342 are almost at right angles (angle $\sim$89.°4).

The perturber would also have a small effect on the motion of the MW, this being $\sim$15 km s^{-1} towards each perturber in the context of this model. For a target near a perturber, one expects them to also be nearby on the sky. Thus, the MW would be pulled towards the target to some extent, reducing its GRV. This might be why our more detailed model for tides (see below) often predicts that they would reduce the GRVs of target galaxies.

Our more detailed model involves two gravitating masses. We treat the MW and M31 as a single object with mass 4×10^{12} M$_\odot$ and assume $q_1 = 0.2$. This object represents the LG. We put the LG and the perturber along the y-axis and solve both objects forwards using equation (32) (the relevant mass is that of the MW+M31+perturber). Their initial separation is varied so as to get a final separation equal to the observed distance between the LG barycentre and the perturber.

We then determine how a target galaxy would fit into this picture. We solve a test particle trajectory so that it ends up at the correct distance from the LG barycentre and at the correct angle to the perturber as perceived at the LG particle.[18] The final GRV of the test particle is determined using equation (50), referred to the LG particle rather than the MW.

To determine the effect of the perturber, we then (effectively) reduce the perturber mass to 0 and repeat the calculation. The final GRV of the test particle is compared between the two simulations. Some results from this procedure are shown in Table 6.

The combination of large distances from the perturbers ($\gtrsim$ 2 Mpc) and projection effects reduce how much tides might have affected the GRVs of target galaxies. As a result, tides from IC 342 and M81 cannot explain the very high HRVs of targets such as the DDO objects in the context of this model. In fact, for several galaxies like these, tides seem to reduce GRVs and thus make the discrepancy even worse. Thus, we do not believe that tides are responsible for the discrepancies, assuming we have reasonable perturber masses (Table 1) and a good method of estimating their effects.

[18] A 2D model is sufficient for this as there are three particles.

Table 6. ΔHRVs for the LG galaxies most discrepant with our model. Error budgets are found by adding $\sigma_{\rm pos}$ and σ_{v_h} in quadrature. We also give an estimate in km s^{-1} for how much M81 and IC 342 might have affected the GRV of each galaxy (method described in text).

Galaxy	ΔHRV (km s^{-1})	Effect of M81	Effect of IC 342
DDO 99	110 ± 13	-6.5	-10.3
DDO 190	104 ± 5	-7.5	-9.8
KKH 98	100 ± 12	-3.0	-9.5
DDO 125	95.5 ± 4.5	-5.9	-10.6
NGC 404	89 ± 37	-3.6	-13.2
Tucana	48 ± 6	2.0	5.4
NGC 3109	38.1 ± 5.1	-1.9	1.4
ESO 294-G010	-16.8 ± 5.8	3.6	4.9
IC 4662	-24 ± 14	3.1	9.4
IC 5152	-32.3 ± 4.0	3.7	7.1
Leo A	-46.8 ± 4.4	-2.3	-3.1

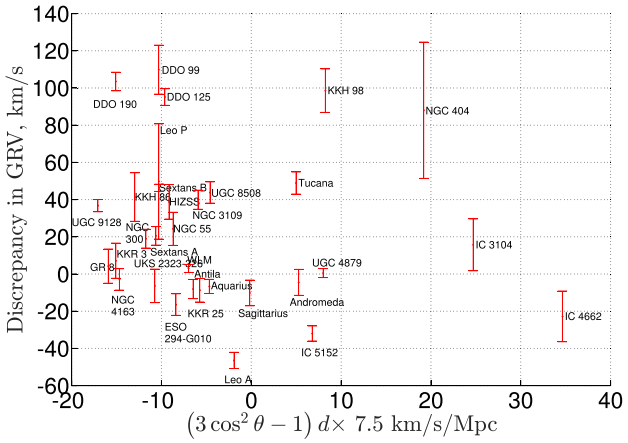


Figure 16. ΔHRVs are plotted against our estimate for how much the GA might have increased the HRV of each galaxy. The distance d used here is heliocentric. Note that the DDO objects would likely have their radial velocities reduced by tides from the GA.

4.3.3 The GA

There are additional structures in the Universe on a larger scale which might be pertinent to our analysis. In particular, the LG as a whole has a velocity of ∼630 km s^{-1} with respect to the surface of last scattering (Kogut et al. 1993). It is thought that this is mostly due to the gravity of the GA (Mieske, Hilker & Infante 2005). Assuming a distance of 84 Mpc, it is clear that tides raised by the GA can have a non-negligible impact on motions within the LG.

As the GA is much more distant from the MW than objects within the LG, we use the distant tide approximation. Treating the MW and other target galaxies as freely falling in the gravitational field of a distant point mass, the change in the GRV of a target galaxy is

$$\Delta\text{GRV}_{\rm GA} = \left(3\cos^2\theta - 1\right)\frac{v_{\rm pec,LG}\, d}{d_{\rm GA}} \quad \text{for } d \ll d_{\rm GA}. \tag{62}$$

Here, $d_{\rm GA}$ is the distance to the GA while θ is the angle on our sky between it and the target galaxy, which is at a heliocentric distance d. The GA is assumed to have caused the LG to gain a peculiar velocity of $v_{\rm pec,\,LG} = 630\,{\rm km\,s^{-1}}$. We take the GA to be in the direction $l = 325^\circ$, $b = -7^\circ$ in Galactic coordinates (Kraan-Korteweg 2000).

For θ close to 0 or 180°, the GA tends to increase GRVs. However, for θ close to 90°, the GA reduces GRVs. This arises because both the MW and the target galaxy fall towards the perturber at similar rates. As their comoving distance from the GA decreases, so also does their comoving distance from each other.

Due to the GA, a test particle started with the same initial conditions will end up with an altered position as well as velocity. Thus, the initial position must be altered to match a fixed final position. This reduces the effect of the GA on predicted velocities of target galaxies (a similar effect is shown in Fig. 4).

Because of Hubble drag, present peculiar velocities are mostly sensitive to tides at late times. Thus, we expect equation (62) to provide a reasonable approximation as long as we have accurate distances to the relevant objects and know their sky positions.

Although we may overestimate the magnitude of $\Delta\text{GRV}_{\rm GA}$, it is much harder to get its sign wrong. This is because the sign is dependent on the factor of $(3\cos^2\theta - 1)$, a quantity sensitive only to the (usually well-known) sky positions of relevant objects but not to their distances. The trajectory of a distant LG galaxy is unlikely to have deviated much from a radial orbit. The GA is much more distant so it was probably always in much the same direction on the sky. As a result, we expect θ to never have been much different to its present value. Although this may not be true at very early times, Hubble drag makes the system 'forget' about the forces acting at such times.

In Fig. 16, we show how much tides from the GA would likely affect the HRV of each target galaxy. It is interesting to see the results for the galaxies with the highest ΔHRVs, in particular the DDO objects (Table 6). Because of their positions on the sky, tides from the GA would actually *reduce* their GRVs. This makes it more difficult to explain their high observed HRVs.

Perhaps more important is the lack of any apparent correlation between HRV discrepancies and the effect of tides raised by the GA. This suggests that it cannot reconcile the differences between our best-fitting simulation and observations. In fact, it would probably make matters worse as it reduces HRVs for three out of the four objects which likely have $\Delta\text{HRV} > 90\,{\rm km\,s^{-1}}$ (i.e. $\Delta\text{HRV} > 3\sigma_{\rm H}$).

To estimate how much the GA might affect $\sigma_{\rm extra}$, we adjusted model-predicted GRVs of all our target galaxies using equation (62). We then re-ran our statistical analysis. This raised $\sigma_{\rm extra}$ by $\sim 6\,{\rm km\,s^{-1}}$ (or $4\,{\rm km\,s^{-1}}$ if the RLG is assumed empty). The actual effect of the GA is probably smaller because it affects final positions of test particles as well as their velocities. Still, it seems likely that the GA makes it harder rather than easier to explain observed HRVs.

Although the only LSS we consider explicitly is the GA, it is of course possible to perform a similar analysis for an external perturber in another direction. Indeed, all possible directions can be investigated using a grid method. This was done by Peñarrubia et al. (2016), who used a 1D model for the LG (see Section 4.7). Fortunately, tides raised by LSS are more important towards the edges of the LG. Here, the greater distance from the MW and M31 makes it more realistic to consider them as a single point mass. Thus, one might expect their analysis to be reasonably sensitive to tides raised by LSS. As a result, it is important to note that they 'found no statistically meaningful deviation between the velocities predicted by the point-mass model and the location of galaxies on the sky'.

4.4 Kinematic corrections due to massive satellites

Massive satellite galaxies of the MW can affect our timing argument analysis because of an indirect kinematic effect. Instead of dealing with just the MW, we should really deal with the MW system (≡ MW + satellite). The brightest satellite galaxy of the MW is the LMC. This is ∼ 50 kpc from the Sun (Pietrzyński et al. 2013). Being much fainter than the MW, we expect it to be much less massive. As a result, it shifts the barycentre of the MW system by $\lesssim 10$ kpc.

Previously, we neglected errors that arise due to the heliocentric directions towards other LG galaxies not being the same as the directions from the centre of the MW. This is because the Sun is only ∼ 8 kpc from there (McMillan 2011). Similarly, we also neglect any errors that arise due to the LMC altering the position of the MW system's barycentre. This is because even the nearest target galaxy is ∼ 800 kpc away. Moreover, the directions from the Sun towards the Galactic Centre and towards the LMC are almost orthogonal, meaning that the errors due to these approximations would add in quadrature rather than linearly.

Unlike the position of the MW system's barycentre, its *velocity* may be significantly altered by the LMC. Consequently, we determined $\boldsymbol{v}_{\rm LMC}$, its space velocity with respect to the MW. This requires knowledge of its HRV (McConnachie 2012) and its proper motion (Kallivayalil et al. 2013) multiplied by its distance (Pietrzyński et al. 2013). The velocity of the Sun with respect to the MW is also required (Table 2).[19] Using these references, we found that the speed of the LMC with respect to the MW is 319 845.6 m s^{-1} of which 229 401.5 m s^{-1} is directed towards the North Galactic Pole. Importantly, the component of this velocity directly away from M31 is 241 223.4 m s^{-1}.

To apply a kinematic correction for the motion of the LMC, we note that the velocity of the Sun with respect to the MW should now be altered to its velocity with respect to the barycentre of the MW system.

$$\boldsymbol{v}_\odot \rightarrow \boldsymbol{v}_\odot - f_{\rm recoil}\, q_{\rm LMC}\, \boldsymbol{v}_{\rm LMC} \tag{63}$$

$$q_{\rm LMC} \equiv \frac{M_{\rm LMC}}{M_{\rm MW} + M_{\rm LMC}}. \tag{64}$$

Here, $\boldsymbol{v}_{\rm LMC}$ is the velocity of the LMC with respect to the MW disc. Note that the MW mass $M_{\rm MW}$ does not include a contribution from the LMC mass $M_{\rm LMC}$. The − sign in equation (63) arises because we are correcting for the recoil induced by the LMC on the MW. If the LMC were bound to the MW and the two were orbiting their common centre of mass, then a simple application of Newton's third law would show that we should set $f_{\rm recoil} = 1$.

It is possible that the LMC is not bound to the MW but rather is on a first infall trajectory (e.g. Besla et al. 2007). Indeed, its high speed relative to the MW means it is unlikely to be a gravitationally bound satellite galaxy (Wu et al. 2008). Although the magnitude of $\boldsymbol{v}_{\rm LMC}$ is now believed to be smaller than the ∼380 km s^{-1} assumed in this work, other considerations continue to suggest that it is unbound (see section 6.4 of Kallivayalil et al. 2013). Thus, most of its velocity might have been present even when it was far from the MW. In this case, only part of its velocity would have been gained due to gravity from the MW. As a result, the recoil of the MW induced by the LMC would be less than if the MW and LMC were bound.

[19] The important quantity here is actually the velocity of the LMC with respect to the Sun, so it is essential to use the same $\boldsymbol{v}_\odot$ as in the rest of our analysis.

To account for this possibility, we introduce the parameter $f_{\rm recoil}$, the fraction of the momentum of the LMC that has been gained due to gravity from the MW. We assume that the direction of the recoil induced by the LMC on the MW is aligned with the MW–LMC relative velocity. The effect of the LMC on our analysis is maximized if we set $f_{\rm recoil} = 1$ and assume the LMC is bound to the MW. The validity of this assumption remains an open question. However, to put an upper limit on the effect of the LMC, we will make this assumption.

In applying a kinematic correction for the LMC, another major uncertainty is its mass relative to the MW. Recent rotation curve measurements of the LMC based on both radial velocities and proper motions indicate a flatline level in its outer regions of ∼ 90 km s^{-1} (Kallivayalil et al. 2013). Extrapolating to a tidal radius of 25 kpc as suggested by this work, we obtain an enclosed mass of $4.7 \times 10^{10}\,{\rm M}_\odot$. The actual value is likely to be smaller because other studies indicate a slower rotating LMC (Alves & Nelson 2000).

The LMC mass can also be estimated using an abundance matching technique. This yields a pre-infall mass of $\sim 1.9 \times 10^{11}\,{\rm M}_\odot$ (Boylan-Kolchin et al. 2010). Not all of this mass can get as close as 50 kpc to the MW and exert a force on it. This is because the outer parts of the LMC's dark matter halo have likely been tidally stripped due to its close approach of the MW. The work of Boylan-Kolchin et al. (2010) suggests that we should reduce the pre-infall mass of the LMC by a factor of ∼ 3.5 to account for this (see their fig. 11). This makes both estimates of the LMC mass agree.

However, several recent investigations suggest a much higher LMC mass, which may be possible if it has not been tidally stripped to a significant extent. This is tied to the issue of whether the LMC is on its first infall into the MW, as first suggested by Besla et al. (2007). Those authors conducted further investigations into this possibility (Besla et al. 2010, 2012). Recently, it was shown that a first infall of a massive LMC could induce a recoil on the MW of as much as ∼70 km s^{-1} (Gómez et al. 2015), corresponding to $q_{\rm LMC} \lesssim 0.2$. A high LMC mass is also hinted at by the discovery of stellar streams around the Magellanic Clouds (Belokurov & Koposov 2016) and by its high star formation rate, suggestive of a first infall (Tollerud et al. 2011).

A very massive LMC would exert strong tides on the disc of the MW, perhaps warping it more than is observed. Assuming a bound LMC, it proved possible to reproduce important properties of the observed warp with a fairly low LMC mass of just $20 \times 10^9\,{\rm M}_\odot$ (Weinberg & Blitz 2006). If the LMC was instead on its first infall, it would only recently have had a substantial effect on the MW. This might be compensated by a higher LMC mass. The interplay between these effects deserves further investigation.

To incorporate the LMC into our analysis, we assumed that the relevant MW mass for the purposes of the timing argument is the combined mass of the MW and the LMC. Even if the LMC was quite far from the MW in the past, it seems likely that other LG galaxies were much further still, so that the MW and LMC can be treated as a single point mass. Neglecting the small increase in MW mass due to accretion (Fig. 6), this means that

$$q_{\rm LMC} = \frac{M_{\rm LMC}}{q_{_1} M_i}. \tag{65}$$

In models with a very low total LG mass M_i and a very small fraction $q_{_1}$ of this in the MW, it is possible to get $q_{\rm LMC} > 1$. To avoid this occurring, we calculated $q_{\rm LMC}$ using equation (65) and then capped its value at 0.3. This is a very generous upper limit on the ratio between the LMC and MW masses – high-resolution

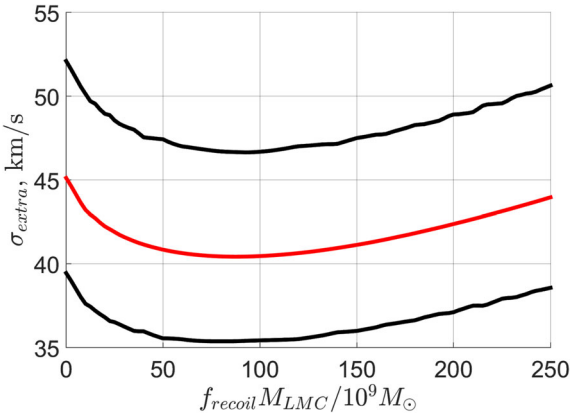


Figure 17. The most likely value of $\sigma_{\rm extra}$ and its 1σ uncertainty are shown here as a function of the LMC mass (assuming $f_{\rm recoil} = 1$, see text). The LMC is included via equation (63). Note that $\sigma_{\rm extra}$ eventually increases with the LMC mass because very high LMC masses imply a very large kinematic correction for it.

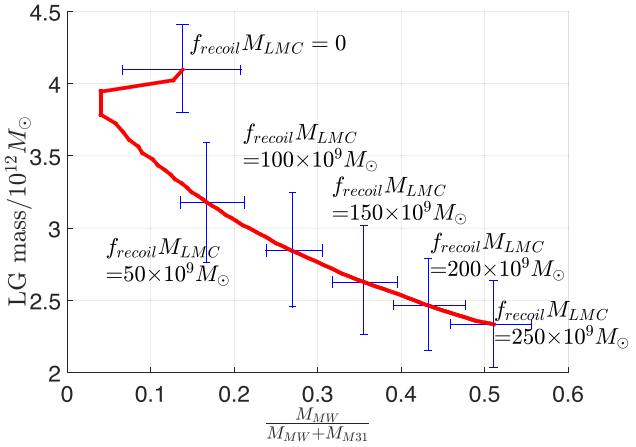


Figure 18. The red line shows the locus of the most likely values of M_i and q_1 as a function of LMC mass. Crosses show uncertainties on each parameter for six different LMC masses. Fig. 7 suggests that uncertainties in M_i and q_1 are nearly uncorrelated.

ΛCDM simulations indicate that it is very unlikely to find a subhalo with >10 per cent as much mass as the main halo (Boylan-Kolchin et al. 2010). Moreover, virial masses scale approximately as the cube of rotation velocities (Evrard et al. 2008). Assuming the MW rotation curve in the DM-dominated regions is above 180 km s^{-1} (Kafle et al. 2012), this suggests an LMC mass of below $\sim\frac{1}{8}$ that of the MW.

To check if our imposed upper limit on $q_{\rm LMC}$ was affecting our analysis, we determined the best-fitting values of M_i and q_1 for each value of $M_{\rm LMC}$ that we tried. We then verified that the most likely total MW system mass ($M_i q_1$) always comfortably exceeded $\frac{10}{3}$ times the LMC mass (in fact, it was never below $4.5 M_{\rm LMC}$ – see Fig. 18). This indicates that our analysis should not have been greatly affected by our decision to cap $q_{\rm LMC}$ at 0.3.

We investigated a wide range of LMC masses (0–250× 10^9 M$_\odot$) and tried 101 linearly spaced values in this range. Each time, we re-calculated the probability density function over the other four model parameters, meaning that we essentially added a fifth parameter using a grid method. As we are only including the kinematic effect of the LMC, it is not necessary to repeat our dynamical simulations. Only the statistical analysis had to be redone. The posterior probability distribution was then marginalized over each of the model parameters one at a time to look for trends with the LMC mass.

Probably the most important result of this investigation is that the overall fit to the observations is not much improved. We quantify this using $\sigma_{\rm extra}$, which is reduced slightly once the LMC is included (Fig. 17). However, for very large LMC masses, the kinematic correction it induces becomes very large, thereby worsening the fit to the data. Thus, including the LMC cannot reduce $\sigma_{\rm extra}$ by even as much as its formal uncertainty.

The correction for the LMC is implemented by altering $\boldsymbol{v}_\odot$ according to equation (63). At given $q_{\rm LMC}$, this adds a constant vector to the predicted velocities of all LG galaxies with respect to the Sun. If we set $q_{\rm LMC} = 0.2$, then the magnitude of this vector is 64 km s^{-1}, comfortably exceeding $\sigma_{\rm extra}$. None the less, including the LMC hardly reduces $\sigma_{\rm extra}$. This is because the galaxies we identified as having anomalously high radial velocities (Table 6) are in several quite different sky directions. Indeed, we confirmed that assuming a large LMC mass causes some galaxies to have HRVs very substantially below the predictions of the best-fitting model.

Table 7. The most likely value of the LSR speed $v_{c,\odot}$ with its 1σ uncertainty is given as a function of the assumed LMC mass. The prior constraint is 239 ± 5 km s^{-1} (McMillan 2011). There is no tension between this value and that suggested by our analysis.

$f_{\rm recoil}\, M_{\rm LMC}$	$v_{c,\odot}$
0	239.53 ± 4.82 km s^{-1}
125× 10^9 M$_\odot$	239.44 ± 4.72 km s^{-1}
250× 10^9 M$_\odot$	$238.55^{+4.65}_{-4.67}$ km s^{-1}

The motion of the LMC with respect to the MW disc is mostly along the MW–M31 line, so one expects a strong effect on the implied total LG mass M_i. This is clearly borne out by our analysis (Fig. 18). The lower LG mass resulting from including the LMC is now more consistent with the works of Peñarrubia et al. (2014) and Diaz et al. (2014) – both give values around 2.5 × 10^{12} M$_\odot$.

Our analysis clearly prefers a non-zero value for $q_{\rm LMC}$ (Fig. 17). Thus, at low LMC masses, the analysis prefers low MW masses to force up $q_{\rm LMC}$ towards its preferred value. The opposite occurs at high LMC masses – a rapid increase in q_1 is required to raise the MW mass and hold down the kinematic correction due to the LMC. This is because the fit to the data is worsened if this correction is too large.

The LMC also affects the tangential velocity of M31 with respect to the MW system. Considering the M31 proper motion measurement of van der Marel et al. (2012a), it is likely that including the LMC slightly increases the tangential velocity of M31, though its radial velocity is increased far more. Still, M31 should remain on a nearly radial orbit with respect to the MW system if the LMC is given a reasonable mass. Any tangential motion would increase the inferred total LG mass as there would be a larger centrifugal force between the MW and M31 (which is not included in our analysis). This would tend to reduce model-predicted GRVs of LG galaxies, making it even more difficult to explain the observations.

Including the LMC hardly affects our inference on $v_{c,\odot}$ (Table 7). We used a prior on this parameter of 239 ± 5 km s^{-1} (McMillan 2011). Our analysis slightly reduces its uncertainty. Based on the magnitude of this reduction, we conclude that our timing argument analysis independently constrains the LSR to within ~ 15 km s^{-1}.

The combination of the prior with our work yields a best-fitting LSR speed very close to that implied by the prior alone. This suggests that if we did not impose a prior constraint on $v_{c,\odot}$, then we would find a most likely value for $v_{c,\odot}$ that was within $\sim \left(\frac{15}{5}\right)^2 \times 0.5 \approx 5$ km s^{-1} of 239 km s^{-1}. Thus, there is no tension between the LSR speed preferred by our analysis and that preferred by McMillan (2011) on independent grounds, although our analysis is ~ 3 times less accurate in this regard.

It is possible for massive satellites of M31 to cause a similar kinematic correction to its adopted HRV [20]. Our analysis places a high statistical weight on M31 by requiring our models to match its HRV fairly well. However, our previous results were almost unaffected if we used the same value of $\sigma_{\rm extra}$ for M31 as for other LG galaxies (the inferred value of $\sigma_{\rm extra}$ differed by ~ 2 km s^{-1}). This remains true if we include the kinematic effect of the LMC in our analysis. In other words, our results are not much different if we treat M31 in exactly the same way as other LG galaxies. Consequently, even a substantial alteration to the HRV of M31 should hardly affect our analysis as it involves 31 other galaxies.

The case is different for the LMC because including a kinematic correction for it alters the predicted HRV of every galaxy in the LG rather than just the observed HRV of one galaxy. As the former does not much affect our overall conclusions, we suspect the same is true of the latter.

4.5 Interactions with massive satellite galaxies

Models invoking gravitational slingshot encounters near the MW/M31 to fling out galaxies at high speed seem to share an important feature with the data: in Fig. 8, most of the galaxies (20/32) are going outwards faster than predicted by the best-fitting model. This trend is perhaps clearer in Fig. 9.

We already included gravity from the MW and M31. One possibility not previously considered is that their satellites could have interacted with what are now non-satellite galaxies in the LG. For example, the Triangulum galaxy (M33) might be able to expand the region around M31 with a disturbed velocity field. This is possible via gravitational slingshot encounters with M33, using energy from its orbital motion around M31 to fling out material at high speed.

Considering that M33 rotates at $\sim$100 km s^{-1} (Corbelli 2003), it cannot have affected the motion of a passing object by much more than this without merging with it. Thus, an important issue with such a scenario is whether it can explain the fast outward motion of galaxies like DDO 125. Not only would it have to reach its present position several Mpc from M31, it would also have to possess sufficient kinetic energy to move at its present high velocity. Even if we neglect the retarding effect of gravity from the MW and M31, Hubble drag alone would mean that a peculiar velocity of 100 km s^{-1} today needed to have been 300 km s^{-1} at redshift 2, a plausible time for the interaction considering how far DDO 125 is from M31 (Fig. 3).

Moreover, one expects only a small fraction of the material in the LG to have interacted with Triangulum in the narrow range of impact parameters that lead to a large impulse but avoid a merger. Some of the material that was unaffected by M33 would no doubt have interacted with the LMC or with other massive satellites. Still, we find it hard to believe that such interactions would be as likely or as strong as required to fit the observations. Achieving both simultaneously does not seem feasible.

Our models did not have particles starting too close to the MW or M31. We mapped the gravitational potential at $t = t_i$ (equation 43) and assumed that all material below a certain level (i.e. with $U < U_{\rm exc}$) had gone into one of these galaxies. For most parameters, this region did not split into separate regions around each galaxy but was a single region encompassing both. Test particles were not allowed to start within it.

It is possible that pockets of material within this 'excluded region' did not get accreted by the MW or M31. Starting closer to one of these galaxies, this material might be more likely to interact with one of their satellites. However, it is unclear how such interactions could have been strong enough to explain the observations as the material would also have a deeper potential well to climb out of.

4.6 Interactions with the MW and M31

Other than the MW and M31, none of the objects in the LG seem heavy enough to impart a sufficiently large impulse on our target galaxies. However, our models already include gravitational slingshots caused by the MW and M31. Such encounters provide a way of extracting energy from the motion of these galaxies and putting it into the motion of a less massive third object.

In principle, Andromeda can exert a large impulse on a passing object – perhaps up to twice Andromeda's rotation speed. Therefore, it might be able to exert an impulse of as much as ~ 450 km s^{-1} on an object which approached closely enough yet avoided merging. For such an interaction to help explain the observations, the scattered object must have been fast-moving relative to M31 *even when the two were far apart*. Otherwise, even fully reversing the small relative velocity 'at infinity' would only lead to a small impulse.

In our simulations, the MW and M31 have never been moving very fast (Fig. 1). Their relative motion has usually been slower than at present ($\sim$110 km s^{-1}; van der Marel et al. 2012a). It is difficult to achieve an impulse much exceeding the motion of the massive body. Supposing the MW was moving at $\sim$90 km s^{-1}, a small fraction of the material in the LG received an impulse of perhaps that much.[21]

The effect of Hubble drag then reduces the peculiar velocity gained in this way. So also does the gravity of M31 (except for particles between the MW and M31, a region in which none of our target galaxies lie). Thus, the region in which the velocity field is disturbed by interactions with the MW/M31 only goes out to $\sim$1 Mpc from the LG barycentre (Fig. 3). At higher LG masses, this region is somewhat larger: its linear size $\propto M_i^{1/3}$. It is difficult to see how this region can be made to encompass the whole LG.

Prior to the start of our simulations, the MW–M31 mutual gravity would not yet have had much time to retard their motion. Thus, we can assume that they were tracing the cosmic expansion, with mutual separation $d(t) \propto a(t)$. At these times, the Hubble parameter $\frac{\dot{a}}{a} \sim a^{-3/2}$ (equation 36) and so we expect the velocities of the galaxies to behave as $a^{-1/2}$. As a result, an interaction with a passing dwarf galaxy could lead to a maximum impulse on it that depends on the encounter time as $\sim a(t)^{-1/2}$. This means that encounters of the MW/M31 with LG dwarfs at very early times may have been very powerful.

[20] Though see Peñarrubia et al. (2016) for why such corrections are likely very small, even for the brightest M31 satellites.

[21] In a logarithmic potential, an extremely eccentric orbit has an angle of $\sim$240° between apocentres, meaning that the deflection angle is $\sim$60° rather than 180°.

However, due to Hubble drag, the effect on the present peculiar velocity of the dwarf galaxy would be reduced by a factor of a at the time of the encounter.[22] This means that very early encounters between the MW/M31 and LG dwarf galaxies should hardly affect our analysis, even if they were very strong. This is probably why our results changed very little when we altered the start time of our simulations to correspond to redshift 14 rather than nine ($\sigma_{\rm extra}$ decreased by $\sim 1\,{\rm km\,s^{-1}}$ when using the earlier start time).

It is unlikely that the MW and M31 existed at earlier times. This makes it difficult to argue that early encounters between LG dwarfs and the MW/M31 are responsible for the anomalously high HRVs of some distant LG galaxies.

Therefore, one possible solution is to suppose that the MW and M31 were moving much faster than in our model a few Gyr after the big bang. We mentioned in Section 1 that they might indeed have done so. In MOND, they would have undergone a close flyby ∼9 Gyr ago (unlike the ΛCDM-based trajectories used in our models, Fig. 1). The relative speed at the time of closest approach would have been $\sim 600\,{\rm km\,s^{-1}}$ (Zhao et al. 2013). One could suppose that the MW was moving at $400\,{\rm km\,s^{-1}}$ and M31 at $200\,{\rm km\,s^{-1}}$. Any passing dwarf galaxies would then have received a large impulse. Hubble drag would reduce peculiar velocities gained in this way by only a factor of ∼3. Thus, the MW–M31 relative speed in this model seems about right to explain the motions of the LG galaxies with high ΔGRVs.

Of course, one does not expect *all* of our target galaxies to have received an impulse quite this large. Only some of the material in the LG would have closely approached the MW or M31 at a time when their relative speed exceeded e.g. $400\,{\rm km\,s^{-1}}$. This material might then get flung outwards and become an observed galaxy. It is also possible for the material to later merge with a galaxy that never strongly interacted with the MW or M31. Depending on the mass ratio of such a merger, it might leave the resulting object with a GRV only a little above that of the unperturbed galaxy before the merger.

An interesting aspect of the observations which may point towards this scenario is apparent in Fig. 10: the galaxies with the greatest excess radial velocity relative to ΛCDM all seem to be towards the edge of the LG. This may be because those objects which were flung out at higher speeds are now further away from the MW and M31.

In this scenario, the high ΔGRV galaxies were all (roughly) at the same place at the same time: close to the LG barycentre when the MW–M31 flyby occurred. Thus, the magnitudes of the GRV discrepancies can be used to estimate the time of the flyby. The highest ΔGRV galaxy in our sample is DDO 99, with $\Delta{\rm GRV} = 100\,{\rm km\,s^{-1}}$ and a distance of ∼3 Mpc.[23] Assuming objects at this distance would nearly follow a pure Hubble flow in ΛCDM (e.g. bottom panel of Fig. 3), its radial velocity should be $H_0 d \sim 200$ km s^{-1} with respect to the LG barycentre. Neglecting projection effects,[24] the actual radial velocity is ∼50 per cent larger. This corresponds to an elapsed time since the flyby of $\sim \frac{2}{3}$ the age of the Universe, i.e. ∼9 Gyr ago.

Interestingly, this is also when there appears to have been a sudden perturbation to the disc of the MW which created its thick disc (Quillen & Garnett 2001). There is some circumstantial evidence that this perturbation was tidal in nature rather than a process internal to the MW (Banik 2014). Although a minor merger is an obvious possibility, it would leave accreted stars. As the accreted galaxy must have been reasonably massive compared to the MW to create its thick disc, these accreted stars should have characteristic properties. Recent attempts to find such stars have not found any (Ruchti et al. 2015). This might be an indication that the thick disc was created by a close flyby of another massive galaxy ∼9 Gyr ago rather than a minor merger at that time.

In this respect, it is interesting to estimate when an MW–M31 interaction might have occurred if MOND were the correct description of nature. Applying this theory, it can be shown that the time from apocentre to pericentre is given by (equations 15 and 29– 30 of Zhao, Li & Bienaymé 2010)

$$\Delta t \approx \int_0^1 \left[\frac{2}{{t_{_M}}^2}\ln\frac{1}{x} - (1-x^2){H_0}^2\Omega_{\Lambda,0}\right]^{-1/2} {\rm d}x \quad \text{where}$$

$$t_{_M} = \frac{d_{\rm apo}}{\sqrt[4]{0.61\,GMa_0}}. \tag{66}$$

Note that the relevant mass M is the total baryonic mass of the LG. The non-linear nature of the theory reduces the force between two particles with the same total mass if it is distributed more equally. Even assuming (conservatively) equal MW and M31 masses (leading to the factor of 0.61) and using a very low estimate for M of $10^{11}\,{\rm M_\odot}$, we would get $\Delta t \approx 7.2$ Gyr.[25] The MW and M31 are slightly past their apocentre now, but it is still clear that they must have had a past close flyby in this theory.

Equation (66) neglects several complications which arise in MOND. Most important is the external field effect (Bekenstein & Milgrom 1984; Milgrom 1986). This arises because MOND is an acceleration-dependent theory. As a result, a constant external gravitational field acting upon a system weakens the self-gravity of objects within the system. This effect is approximately taken into account in the work of Zhao et al. (2013).

We are currently undertaking more accurate timing argument calculations in the context of MOND. Preliminary results indicate that the LG mass implied by the MOND timing argument is consistent with baryons only, if one requires a past close approach between the galaxies. This flyby needs to have been ∼8–9 Gyr ago. It is not feasible to construct trajectories of the MW and M31 in MOND which avoid such a close approach or have ≥2 such events.

In MOND, the longer range nature of gravity means that we need a more careful treatment of objects outside the LG (Table 1). It is possible that some of the anomalously high outwards velocities found in this work are due to LG galaxies – especially those close to perturbers – falling in towards them. Indeed, including tides from Centaurus A improved the fit to observations somewhat, even in the context of ΛCDM (Fig. 14). This effect might be further enhanced in modified gravity scenarios.

Interestingly, the discrepancy does seem to be higher for target galaxies closer to M81 or to IC 342 (Fig. 15). We argued that the apparent correlation could not be due to gravity from these perturbers as this would imply that they affected velocities of target

[22] Fig. 4 suggests that a factor of $a^{2.4}$ might be more accurate.

[23] As this is the highest ΔGRV out of 32 galaxies, the true value is likely smaller than the $110\,{\rm km\,s^{-1}}$ obtained using nominal values.

[24] On the sky, DDO 99 and M31 are almost at right angles (Fig. 3), so our conclusions should not be much affected by uncertainty in the motion of the MW due to uncertainty in q_1.

[25] We used cosmological parameters as in Table 2 and an apocentre distance $d_{\rm apo} = 1$ Mpc, slightly less than in ΛCDM due to the stronger gravity. For a_0, we used $1.2 \times 10^{-10}\,{\rm m\,s^{-2}}$ (McGaugh 2011).

galaxies more than is reasonable. However, this argument likely breaks down under a different law of gravity. In this case, tides might well be more significant than we assumed.

4.7 Comparison with Peñarrubia et al. (2014)

Our results are broadly similar to the recent study conducted by Peñarrubia et al. (2014, hereafter P14). Those authors also favour a low q_1 and a similar LG mass. We found a slightly higher value of $\sigma_{\rm extra}$ ($45^{+7}_{-6}\,{\rm km\,s^{-1}}$ instead of $35^{+6}_{-4}\,{\rm km\,s^{-1}}$), though the estimates are marginally in agreement.

However, our investigation greatly strengthens the conclusions reached by P14. We used an axisymmetric model for the LG rather than a spherically symmetric one. As our target galaxies are often not much further away than the MW–M31 separation, a spherically symmetric gravitational field may be a poor approximation. For example, gravitational slingshot encounters with the MW/M31 rely on a time-dependent non-spherical potential. Without modelling either of these effects, it would be difficult to draw reliable conclusions about whether these close encounters might have left an imprint on the present motions of target galaxies.

Even if one could be sure that such encounters were not important, a two-centred potential has other subtle consequences. A trajectory which initially went orthogonal to the MW–M31 line from the point halfway between them; would curve towards the heavier galaxy (almost certainly M31). This would tend to increase the HRV of the target galaxy (e.g. bottom panel of Fig. 5). Curvature of test particle trajectories seems to be important even at quite large distances from the MW and M31 (top panel of Fig. 3). The process is more significant if the MW and M31 masses are very unequal, which definitely seems to be the case (Fig. 7).

By using the same list of target galaxies as P14, we avoid targets too close to any of the major perturbers relevant to our analysis (Table 1). However, tides from these objects must affect our results at some level. We directly include the most massive perturber (Centaurus A), exploiting its location almost along the MW–M31 line (Section 2.2.2). We think this greatly improves our model. We conduct a more thorough investigation of tides raised by other objects (Section 4.3) and consider the effect of the LMC in some detail (Section 4.4).

Our initial conditions are handled differently to P14. We use cosmological initial conditions (equation 37) because of observations indicating very low peculiar velocities at early times (Planck Collaboration XIII 2015). P14 used a procedure involving non-cosmological initial conditions which does not seem entirely physical (see their section 3).

We added an extra term to our equation of motion (equation 40) to account for cosmological expansion. The idea is to recover $\boldsymbol{r} \propto a\,(t)$ in the absence of inhomogeneities. A similar approach was used by P14. However, they did not include the deceleration to the expansion rate caused by matter, leaving only the acceleration caused by dark energy. This implicitly assumes that the RLG is empty. As we did not make this assumption, we suspect that the predicted HRVs in our investigation are lower. With a present dark energy fraction of ~ 0.7, we expect a difference of $\sim \left(1-\sqrt{0.7}\right) H_0 d$ for a target at distance d. Assuming a typical distance of 2 Mpc, this suggests a difference of $\sim$20 km s^{-1}. Because observed HRVs tend to exceed predicted ones, it is unsurprising that our estimate of $\sigma_{\rm extra}$ is higher.

In Section 4.1, we discuss how much mass might actually be in the RLG. Here, we show that, although it is possible to have an empty RLG, this is an extreme case. It is also possible for it to have even more mass than we assumed. Our assumption is an intermediate case. None the less, we performed calculations assuming an empty RLG and showed that this reduces $\sigma_{\rm extra}$ by $\sim 7\,{\rm km\,s^{-1}}$.

We argued that M31 should be treated specially in that one should use a lower value of $\sigma_{\rm extra}$ for it than for other LG galaxies. This forces our models to match the HRV of M31 very well, restricting which models are viable and thus raising $\sigma_{\rm extra}$. The effect is substantial for our analysis without Cen A: the most likely value of $\sigma_{\rm extra}$ is raised from 46 to 54 km s^{-1}. However, in our more realistic analysis including Cen A, $\sigma_{\rm extra}$ is only affected by $\sim 2\,{\rm km\,s^{-1}}$. Thus, although we recommend treating M31 specially due to its much higher mass than LG dwarf galaxies, our overall conclusions are little altered if one does not do so. The parameter most affected seems to be M_i, which is lower if M31 is not treated specially. This is also apparent in fig. 13 of P14, especially if one imposes an independent constraint on the LSR speed (e.g. McMillan 2011).

5 CONCLUSIONS

We performed a careful dynamical analysis of the LG to try and explain the observed positions and velocities of galaxies within it. The LG was treated as a collection of test particles and two massive ones – the MW and Andromeda (M31) – which we put on a radial orbit. We added a third massive particle to represent Centaurus A, which is very close to the MW–M31 line (Table 1). All particles were started moving outwards from the centre of mass of the LG with speeds proportional to distance from there. Thus, they all started on a pure Hubble flow with no peculiar velocity (equation 37).

A wide range of possible masses for the MW and M31 was investigated using a grid method (Table 2). Each time, we got the final MW–M31 and MW–Cen A distances to match their observed values. We also got a test particle trajectory to end at the same location as each observed LG galaxy. This gave a model-predicted velocity, whose line-of-sight component (the HRV) was compared with observations.

The best-fitting total LG mass is $4.33^{+0.37}_{-0.32} \times 10^{12}\,{\rm M}_\odot$, with 0.14 ± 0.07 of this accounted for by the MW and the rest by M31. There is almost no tension between the LSR rotation speed estimated by McMillan (2011) and our analysis (Table 7).

However, even in the best-fitting model, there was a poor match between observed and model-predicted HRVs. Thus, we tried to quantify the extra astrophysical noise $\sigma_{\rm extra}$ that the observations imply. To do this, we added it in quadrature with the other known sources of error, which are all observational. $\sigma_{\rm extra}$ can be constrained using

$$P({\rm Model}\,|\,{\rm HRV_{obs}}) \propto \frac{1}{\sigma} {\rm e}^{-1/2\left(\frac{{\rm HRV_{obs}}-{\rm HRV_{model}}}{\sigma}\right)^2}. \quad (67)$$

Raising $\sigma_{\rm extra}$ – and thus σ – will eventually cause the probability to decrease[26]. In this way, we found that $\sigma_{\rm extra} = 45.1^{+7.0}_{-5.7}$ km s^{-1} (Fig. 7). This rather high value is partly due to some galaxies receding from the LG even faster than a pure Hubble flow (Fig. 11). This is despite the attractive gravity of the MW and M31.

We expect interactions between LG dwarf galaxies to have contributed somewhat to $\sigma_{\rm extra}$, but only at the $\sim$15 km s^{-1} level. This is because they have fairly low rotation velocities/internal velocity dispersions (Kirby et al. 2014), limiting how much they can influence each other gravitationally. If we consider just those galaxies with HRVs below the predictions of the best-fitting model (blue

[26] When $\sigma > |{\rm HRV_{obs}} - {\rm HRV_{model}}|$.

area in Fig. 9), then we see that they can indeed be described quite well by a Gaussian of this width.

One might expect the same to be true for LG galaxies with HRVs that exceed the model predictions. However, this is not true (red area in Fig. 9). Thus, there is a systematic trend for radial velocities to be higher than we expect.

We considered tides from objects outside the LG at some length (Section 4.3). The most relevant perturbers are given in Table 1. A correlation is apparent whereby the most problematic galaxies are closest to these perturbers (Fig. 15). Thus, we constructed a simplified model to estimate how much they might have affected the GRVs of the most discrepant galaxies. Due to a combination of projection effects and large distances from the perturbers ($\gtrsim 2$ Mpc), we consider it unlikely that tides from IC 342 and M81 can reconcile our model with observations (Table 6). In fact, they would likely exacerbate the tension in several cases. This also seems to be true of tides raised by the GA (Fig. 16): including these raises $\sigma_{\rm extra}$ by $\sim 5\,{\rm km\,s^{-1}}$ (Section 4.3.3).

A local underdensity may help to explain the observations. We consider this possibility in Section 4.1 and show that this can reduce $\sigma_{\rm extra}$ by at most $\sim 7\,{\rm km\,s^{-1}}$. Increasing the Hubble constant H_0 has a similar effect, which we consider in Section 4.2. This reduces $\sigma_{\rm extra}$ by $\sim 5\,{\rm km\,s^{-1}}$ (Fig. 14). Although H_0 is known quite accurately (Planck Collaboration XIII 2015), a higher value may be appropriate for the LG if there is a local underdensity.

Satellite galaxies of the MW can affect our analysis kinematically. In this regard, we considered the LMC in some detail (Section 4.5). Including the LMC can reduce $\sigma_{\rm extra}$ by at most $5\,{\rm km\,s^{-1}}$ (Fig. 17). Thus, incorporating it into our models does not reconcile them with observations.

Interactions of LG dwarf galaxies with massive MW/M31 satellites (e.g. M33) would probably be too weak to explain distant non-satellite galaxies moving outwards even faster than a pure Hubble flow. As for encounters with the MW and M31 themselves, this would naturally explain why LG galaxies tend to move outwards much faster than expected. However, this process is already included in our simulations – it seems to be too weak.

It might have been more powerful in reality if the MW and M31 had been moving much faster than in our models (Section 4.6). Although this would be very unusual in the context of ΛCDM, such fast motions arise naturally in certain modified gravity theories (Llinares, Zhao & Knebe 2009). For example, MOND (Milgrom 1983) leads to a past close encounter between the MW and M31 and a maximum relative speed of $\sim 600\,{\rm km\,s^{-1}}$ (Zhao et al. 2013).

One aspect of our results which may point towards this scenario lies in the distances to the objects with the highest radial velocities relative to the best-fitting ΛCDM model. These distances all seem to be close to the upper limit of the distance range probed by our sample (Fig. 10). This may be because those objects which received the strongest gravitational boost from a close encounter with the MW/M31 are currently furthest away from the location of the encounter. Considering the distances and velocities of such galaxies suggests that the encounter was ∼9 Gyr ago. This is around the same age as the thick disc of the MW (Quillen & Garnett 2001). It is also roughly when a MOND timing argument calculation suggests that an MW–M31 encounter took place (Zhao et al. 2013).

We think it is likely that the observed positions and velocities of LG galaxies would be easier to explain if there was a past close MW–M31 flyby. This is possible only within the framework of a modified gravity theory. More work will be required to test such a scenario.

ACKNOWLEDGEMENTS

The authors wish to thank Jorge Peñarubbia for helpful discussions and the referee for their patience while we implemented their suggestions, which greatly strengthened this contribution. IB is supported by a Science and Technology Facilities Council studentship. The algorithms were set up using MATLAB®.

REFERENCES

Alves D. R., Nelson C. A., 2000, ApJ, 542, 789
Angus G. W., Diaferio A., Kroupa P., 2011, MNRAS, 416, 1401
Aragon-Calvo M. A., Silk J., Szalay A. S., 2011, MNRAS, 415, L16
Banik I., 2014, preprint (arXiv:1406.4538)
Barnes J. E., Hernquist L., 1992, Nature, 360, 715
Begum A., Chengalur J. N., Karachentsev I. D., Sharina M. E., 2005, MNRAS, 359, L53
Bekenstein J., Milgrom M., 1984, ApJ, 286, 7
Belokurov V., Koposov S. E., 2016, MNRAS, 456, 602
Besla G., Kallivayalil N., Hernquist L., Robertson B., Cox T. J., van der Marel R. P., Alcock C., 2007, ApJ, 668, 949
Besla G., Kallivayalil N., Hernquist L., van der Marel R. P., Cox T. J., Kereš D., 2010, ApJ, 721, L97
Besla G., Kallivayalil N., Hernquist L., van der Marel R. P., Cox T. J., Kereš D., 2012, MNRAS, 421, 2109
Bovy J., Rix H.-W., 2013, ApJ, 779, 115
Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., 2010, MNRAS, 406, 896
Carignan C., Chemin L., Huchtmeier W. K., Lockman F. J., 2006, ApJ, 641, L109
Cautun M., Bose S., Frenk C. S., Guo Q., Han J., Hellwing W. A., Sawala T., Wang W., 2015, MNRAS, 452, 3838
Collins M. L. M. et al., 2015, ApJ, 799, L13
Corbelli E., 2003, MNRAS, 342, 199
Courteau S., Widrow L. M., McDonald M., Guhathakurta P., Gilbert K. M., Zhu Y., Beaton R. L., Majewski S. R., 2011, ApJ, 739, 20
Crain R. A. et al., 2015, MNRAS, 450, 1937
Croxall K. V., van Zee L., Lee H., Skillman E. D., Lee J. C., Côté S., Kennicutt R. C., Jr, Miller B. W., 2009, ApJ, 705, 723
Diaz J. D., Koposov S. E., Irwin M., Belokurov V., Evans N. W., 2014, MNRAS, 443, 1688
Evrard A. E. et al., 2008, ApJ, 672, 122
Famaey B., McGaugh S. S., 2012, Living Rev. Relativ., 15, 10
Fardal M. A. et al., 2013, MNRAS, 434, 2779
Francis C., Anderson E., 2014, Celest. Mech. Dyn. Astron., 118, 399
Gerke J. R., Kochanek C. S., Prieto J. L., Stanek K. Z., Macri L. M., 2011, ApJ, 743, 176
Gibbons S. L. J., Belokurov V., Evans N. W., 2014, MNRAS, 445, 3788
Gillet N., Ocvirk P., Aubert D., Knebe A., Libeskind N., Yepes G., Gottlöber S., Hoffman Y., 2015, ApJ, 800, 34
Gómez F. A., Besla G., Carpintero D. D., Villalobos Á., O'Shea B. W., Bell E. F., 2015, ApJ, 802, 128
González R. E., Kravtsov A. V., Gnedin N. Y., 2014, ApJ, 793, 91
Harris G. L. H., Rejkuba M., Harris W. E., 2010, PASA, 27, 457
Ibata R. A. et al., 2013, Nature, 493, 62
Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2012, ApJ, 761, 98
Kahn F. D., Woltjer L., 1959, ApJ, 130, 705
Kallivayalil N., van der Marel R. P., Besla G., Anderson J., Alcock C., 2013, ApJ, 764, 161
Karachentsev I. D., 2005, AJ, 129, 178
Karachentsev I. D., Kashibadze O. G., 2006, Astrophysics, 49, 3
Karachentsev I. D. et al., 2007, AJ, 133, 504
Kirby E. N., Bullock J. S., Boylan-Kolchin M., Kaplinghat M., Cohen J. G., 2014, MNRAS, 439, 1015
Kogut A. et al., 1993, ApJ, 419, 1

Kraan-Korteweg R. C., 2000, in Page D., Hirsch J. G., eds, Lecture Notes in Physics, Vol. 556, From the Sun to the Great Attractor. Springer-Verlag, Berlin, p. 301
Kroupa P., Theis C., Boily C. M., 2005, A&A, 431, 517
Lehner N., Howk J. C., Wakker B. P., 2015, ApJ, 804, 79
Li Y.-S., White S. D. M., 2008, MNRAS, 384, 1459
Llinares C., Zhao H. S., Knebe A., 2009, ApJ, 695, L145
López-Corredoira M., Kroupa P., 2016, ApJ, 817, 75
Lynden-Bell D., 1981, The Observatory, 101, 111
McConnachie A. W., 2012, AJ, 144, 4
McConnachie A. W. et al., 2009, Nature, 461, 66
McGaugh S. S., 2011, Physical Review Letters, 106, 121303
McGaugh S., Milgrom M., 2013, ApJ, 775, 139
McGaugh S. S., Wolf J., 2010, ApJ, 722, 248
McMillan P. J., 2011, MNRAS, 414, 2446
Majewski S. R., Skrutskie M. F., Weinberg M. D., Ostheimer J. C., 2003, ApJ, 599, 1082
Mieske S., Hilker M., Infante L., 2005, A&A, 438, 103
Milgrom M., 1983, ApJ, 270, 365
Milgrom M., 1986, ApJ, 302, 617
Mirabel I. F., Dottori H., Lutz D., 1992, A&A, 256, L19
Newberg H. J. et al., 2002, ApJ, 569, 245
Partridge C., Lahav O., Hoffman Y., 2013, MNRAS, 436, L45
Pawlowski M. S., 2016, MNRAS, 456, 448
Pawlowski M. S., Kroupa P., 2013, MNRAS, 435, 2116
Pawlowski M. S., McGaugh S. S., 2014a, MNRAS, 440, 908
Pawlowski M. S., McGaugh S. S., 2014b, ApJ, 789, L24
Pawlowski M. S. et al., 2014, MNRAS, 442, 2362
Pawlowski M. S., Famaey B., Merritt D., Kroupa P., 2015, ApJ, 815, 19
Peñarrubia J., Ma Y.-Z., Walker M. G., McConnachie A., 2014, MNRAS, 443, 2204 (P14)
Peñarrubia J., Gómez F. A., Besla G., Erkal D., Ma Y.-Z., 2016, MNRAS, 456, L54
Pietrzyński G. et al., 2013, Nature, 495, 76
Planck Collaboration XIII 2015, preprint (arXiv:1502.01589)
Quillen A. C., Garnett D. R., 2001, in Funes J. G., Corsini E. M., eds, ASP Conf. Ser. Vol. 230, Galaxy Disks and Disk Galaxies. Astron. Soc. Pac., San Francisco, p. 87
Recchi S., Kroupa P., Ploeckinger S., 2015, MNRAS, 450, 2367
Rigault M. et al., 2015, ApJ, 802, 20
Ruchti G. R. et al., 2015, MNRAS, 450, 2874
Salem M., Besla G., Bryan G., Putman M., van der Marel R. P., Tonnesen S., 2015, ApJ, 815, 77
Salomon J.-B., Ibata R. A., Famaey B., Martin N. F., Lewis G. F., 2016, MNRAS, 456, 4432
Sandage A., 1986, ApJ, 307, 1
Sawala T. et al., 2014, preprint (arXiv:1412.2748)
Sawala T. et al., 2016, MNRAS, 457, 1931
Schaye J. et al., 2015, MNRAS, 446, 521
Schmidt K. H., 1958, Astron. Nachr., 284, 76
Schönrich R., 2012, MNRAS, 427, 274
Slipher V. M., 1913, Lowell Obs. Bull., 2, 56
Tollerud E. J., Boylan-Kolchin M., Barton E. J., Bullock J. S., Trinh C. Q., 2011, ApJ, 738, 102
van der Marel R. P., Fardal M., Besla G., Beaton R. L., Sohn S. T., Anderson J., Brown T., Guhathakurta P., 2012a, ApJ, 753, 8
van der Marel R. P., Besla G., Cox T. J., Sohn S. T., Anderson J., 2012b, ApJ, 753, 9
Veljanoski J. et al., 2014, MNRAS, 442, 2929
Weinberg M. D., Blitz L., 2006, ApJ, 641, L33
Wetzstein M., Naab T., Burkert A., 2007, MNRAS, 375, 805
Wojtak R., Knebe A., Watson W. A., Iliev I. T., Heß S., Rapetti D., Yepes G., Gottlöber S., 2014, MNRAS, 438, 1805
Wu X., Famaey B., Gentile G., Perets H., Zhao H., 2008, MNRAS, 386, 2199
Wu P.-F., Tully R. B., Rizzi L., Dolphin A. E., Jacobs B. A., Karachentsev I. D., 2014, AJ, 148, 7
York D. G. et al., 2000, AJ, 120, 1579
Zhao H., Li B., Bienaymé O., 2010, Phys. Rev. D, 82, 103001
Zhao H., Famaey B., Lüghausen F., Kroupa P., 2013, A&A, 557, L3

This paper has been typeset from a TEX/LATEX file prepared by the author.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **467,** 2180–2198 (2017)
Advance Access publication 2017 January 22

doi:10.1093/mnras/stx151

Dynamical history of the Local Group in ΛCDM – II. Including external perturbers in 3D

Indranil Banik★ and Hongsheng Zhao

Scottish Universities Physics Alliance, University of St Andrews, North Haugh, St Andrews, Fife KY16 9SS, UK

Accepted 2017 January 17. Received 2016 December 18; in original form 2016 September 11

ABSTRACT

We attempt to fit the observed radial velocities (RVs) of ∼30 Local Group (LG) galaxies using a 3D dynamical model of it and its immediate environment within the context of the standard cosmological paradigm, Λ cold dark matter (CDM). This extends and confirms the basic results of our previous axisymmetric investigation of the LG. We find that there remains a tendency for observed RVs to exceed those predicted by our best-fitting model. The typical mismatch is slightly higher than in our 2D model, with a root mean square value of ∼50 km s^{-1}. Our main finding is that including the 3D distribution of massive perturbing dark matter haloes is unlikely to help greatly with the high-velocity galaxy problem. None the less, the 2D and 3D results differ in several other ways such as which galaxies' RVs are most problematic and the preferred values of parameters common to both models. The anomalously high RVs of several LG dwarfs may be better explained if the Milky Way (MW) and Andromeda (M31) were once moving much faster than in our models. This would allow LG dwarfs to gain very high RVs via gravitational slingshot encounters with a massive fast-moving galaxy. Such a scenario is possible in some modified gravity theories, especially those that require the MW and M31 to have previously undergone a close flyby. In a ΛCDM context, however, this scenario is not feasible as the resulting dynamical friction would cause a rapid merger.

Key words: methods: data analysis – methods: numerical – Galaxy: kinematics and dynamics – galaxies: groups: individual: Local Group – cosmological parameters – dark matter.

1 INTRODUCTION

The dynamics of the Local Group (LG) of galaxies provided an early indication that our current understanding of physics is insufficient to explain the dynamics of astrophysical systems. Although the Universe must have started off expanding, the Andromeda (M31) and Milky Way (MW) galaxies are currently approaching each other at ∼110 km s^{-1} (Slipher 1913; Schmidt 1958). Their initial recession could not have been turned around in the ∼14 Gyr (Planck Collaboration XIII 2016) since the big bang if the luminous masses of these galaxies attract each other according to the inverse square law of Newtonian gravity (Kahn & Woltjer 1959).

The most commonly accepted solution is that most galaxies – including the MW and the M31 – are surrounded by large amounts of dark matter (e.g. Ostriker & Peebles 1973). For a while, it was thought that this could be non-luminous conventional matter such as stellar remnants (e.g. Carr 1994, and references therein). However, gravitational microlensing searches for such massive compact halo objects indicated that there was not enough mass in them (Alcock et al. 2000; Tisserand et al. 2007). Thus, the required dark matter is thought to consist of an undiscovered stable particle, or at least one with a decay time longer than the age of the Universe (e.g. Steigman & Turner 1985, and references therein). Multidecade searches for this particle have now ruled out a substantial part of the parameter space that was thought to be feasible before the searches started (e.g. Fermi-LAT Collaboration 2015; LUX Collaboration 2017; PandaX-II Collaboration 2016).

Knowing only the separation and relative velocity of two galaxies, it would be very difficult to rule out this scenario. These two pieces of information are sufficient to constrain the relevant model parameters: the initial MW–M31 comoving separation and their combined mass, some of which would lie beyond their virial radii (Peñarrubia & Fattahi 2016).[1] Fortunately, much additional information has recently become available in the form of positions and velocities of many other LG galaxies (e.g. McConnachie 2012, and references therein).

The velocity field traced out by these galaxies should be understandable using the same MW and M31 total mass as is required to explain their present relative motion. An early attempt at such an

★ E-mail: ib45@st-andrews.ac.uk

[1] We assume throughout that the masses of all galaxies do not change over time.

analysis was made by Sandage (1986). There was some difficulty in matching all the data then available.

A more recent analysis based primarily on the catalogue of McConnachie (2012) also treated the gravitational field of the LG as spherically symmetric (Peñarrubia et al. 2014). Later, an adjustment was made for the effect of the Large Magellanic Cloud (LMC) on the MW and thus on the observed velocities of all galaxies in the LG (Peñarrubia et al. 2016).

The MW and M31 can only be treated as a single point mass if their mutual separation d (783 ± 25 kpc; McConnachie 2012) is much less than the distance to the galaxy one is interested in. However, the LG only extends out to ∼3 Mpc, making this assumption not very accurate within it. Further away, other massive objects besides the MW and M31 must also be considered.

Peñarrubia et al. (2014) considered the effects that the MW–M31 quadrupole might have to get the lowest order correction to the spherical symmetry assumption in their work (see their section 2.4). However, at distances from the LG barycentre of only ∼2 × the MW–M31 separation, it is likely that higher order terms would also be important. Moreover, the quadrupole term was not rigorously included in their final analysis.

Thus, we constructed an axisymmetric model of the LG in ΛCDM. We consider this reasonable because the low proper motion of M31 suggests an almost radial MW–M31 orbit (van der Marel et al. 2012a). A major nearby perturber to the LG is the Centaurus A group of galaxies (Harris, Rejkuba & Harris 2010). This lies very close to the MW–M31 line, allowing us to incorporate it into our model. Therefore, our simulation included three massive objects, with LG dwarfs treated as test particles.

We previously published results based on this 2D model (Banik & Zhao 2016) and review it here (Section 2). Motivated by a poor match between the model and observations, we consider a 3D model of the LG (Section 3). The results obtained using it are described in Section 4. We go on to discuss how they compare with those obtained using our 2D model (Section 5). Here, we also consider a few factors beyond those directly included in our models. Section 5.1 is devoted to the effects of the Great Attractor and the Virgo Cluster (VC) on the LG. In Section 5.2, we discuss the possible effects of a departure from the Newtonian gravity law assumed elsewhere in this work. Our conclusions are provided in Section 6.

2 REVIEW OF 2D AXISYMMETRIC MODEL

2.1 Governing equations

We begin by reviewing our axisymmetric dynamical model of the LG (Banik & Zhao 2016), which in turn follows on from the earlier spherically symmetric analysis of Peñarrubia et al. (2014). Our simulations start at a redshift of 9, when the expansion of the Universe was nearly homogeneous (Planck Collaboration XIII 2016). Thus, we assume that everything was following a smooth Hubble flow at that time. This means that the velocity $\boldsymbol{v}$ of each simulated particle would depend on its position $\boldsymbol{r}$ according to

$$\boldsymbol{v}_i = H_i \boldsymbol{r}_i \tag{1}$$

For any quantity k, we use k_i to denote its value at the time when our simulations are started and $\dot{k}$ to denote its time derivative. The expansion rate of the Universe is quantified by the Hubble parameter $H \equiv \frac{\dot{a}}{a}$, where a is the cosmic scale-factor. At the present time, $H = H_0$ and $a = 1$. In a flat Universe containing only matter and dark energy, their values at other times are given implicitly by

$$H(t) = H_0 \sqrt{\frac{\Omega_{m,0}}{a^3(t)} + \Omega_{\Lambda,0}} \tag{2}$$

We use a standard flat ($\Omega_{m,0} + \Omega_{\Lambda,0} \equiv 1$) dark energy-dominated cosmology whose parameters are given at the bottom of Table 2. Defining time t to start at the big bang ($a = 0$) and imposing the current expansion rate of the Universe as a boundary condition, we get that

$$a(t) = \left(\frac{\Omega_{m,0}}{\Omega_{\Lambda,0}}\right)^{\frac{1}{3}} \sinh^{\frac{2}{3}}\left(\frac{3}{2}\sqrt{\Omega_{\Lambda,0}}\, H_0 t\right) \tag{3}$$

LG dwarf galaxies are represented as test particles affected by the expansion of the Universe and by three massive particles – the MW, M31 and Centaurus A. The dynamics of test particles in such situations can be understood using General Relativity (Banik & Zhao 2016, section 2.1).

We constrain the massive particles to move along a line, making our model axisymmetric. Starting with a plane-polar grid of initial positions, we advance the trajectories of a large number of test particles using the equation of motion

$$\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r} - \sum_{\substack{\text{j=MW,}\\ \text{M31,CenA}}} \frac{GM_j\left(\boldsymbol{r} - \boldsymbol{r}_j\right)}{\left(|\boldsymbol{r} - \boldsymbol{r}_j|^2 + {r_{S,j}}^2\right)^{\frac{1}{2}} |\boldsymbol{r} - \boldsymbol{r}_j|^2} \tag{4}$$

r_S is chosen so that the force at $r \ll r_S$ leads to the correct flatline level of rotation curve for each galaxy, i.e. $r_S = \frac{GM}{{v_f}^2}$ where M is the mass of the relevant galaxy, whose rotation curve flatlines at the level v_f. For the MW, we take $v_f = 180$ km s^{-1} (Kafle et al. 2012) while for Andromeda, we use $v_f = 225$ km s^{-1} (Carignan et al. 2006). Fixing v_f meant that we had to adjust r_S for the MW and M31 depending on their respective masses. Because separations between the massive galaxies are always quite large, for simplicity we use a pure inverse square law for the forces between them (i.e. r_S is set to 0 when calculating forces between the three massive particles). For test particles that get within ∼15 kpc of a massive galaxy,[2] we simply terminate the trajectory. It is likely that any real LG dwarf in this situation would be severely disrupted. Moreover, our analysis is concerned with LG dwarfs much further from any of the three massive galaxies in our model (Fig. 1).

Our algorithm advances trajectories using a fourth-order Runge–Kutta method based on an adaptive but quantized timestep, ensuring that the positions of the massive particles are available when needed. The timestep is adapted based on the distances between the object being advanced and the massive galaxies that influence its motion.

Our model is designed to accurately match the presently observed positions of all galaxies within the LG as well as Centaurus A. Thus, we used a 2D Newton–Raphson algorithm to vary the initial relative positions of the MW, M31 and Cen A along a line in order to match their presently observed configuration to within an accuracy of ∼10^{-4}. Solutions involving collisions between any of these galaxies were of course discarded. We were able to obtain a valid solution in all cases, though this was much easier if the algorithm was slightly under-relaxed for better stability.

Once the trajectories of the massive objects were known, we used equation (4) to advance test particle trajectories. The resulting velocity field in one of our simulations is shown in Fig. 1.

[2] 31 kpc for Centaurus A.

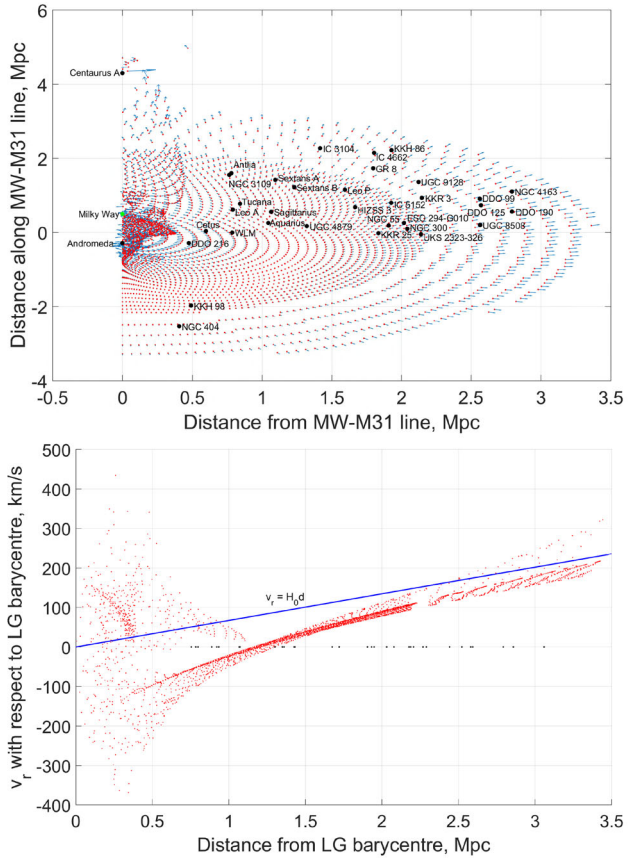


Figure 1. Top: LG velocity field for our best-fitting axisymmetric simulation (parameters given in Table 2). The radial resolution was degraded beyond a distance of 2.3 Mpc as the velocity field is generally quite smooth there. Locations of indicated galaxies are shown relative to the MW–M31 line. Bottom: RVs of test particles with respect to the LG barycentre. Black dots on the x-axis show the distances of target galaxies from there. Without proper motions, observations cannot be put on such a Hubble diagram because the MW is not at the LG barycentre.

To get a test particle landing very close to the observed position of each LG dwarf galaxy, we could simply use a very dense grid of initial positions for our test particles. However, this would be very computationally intensive, especially if extended to three dimensions. Thus, we used a modest resolution grid and found which test particle landed closest to each target galaxy. We then used a 2D Newton–Raphson algorithm to vary the initial position of this test particle, targeting the presently observed position of the corresponding galaxy. Because varying the trajectory of a test particle does not alter the gravitational field in the LG, we were able to improve the accuracy slightly to $\sim 10^{-5}$.

The present velocity of the test particle on this trajectory is our model prediction for the velocity of the target galaxy it represents. We subtract the simulated velocity of the MW and then project the relative velocity on to the direction towards the target to get its model-predicted Galactocentric radial velocity (GRV).

To convert observed heliocentric radial velocities (HRVs) into Galactocentric ones, we need an independent constraint on $\boldsymbol{v}_\odot$, the motion of the Sun within the MW. Part of the challenge is determining $v_{c,\odot}$, the speed of a test particle on a circular orbit around the MW at the position of the Sun. This is called the Local Standard of Rest (LSR). The other part is the non-circular component of $\boldsymbol{v}_\odot$, which consists of $U_\odot$ towards the Galactic Centre, $V_\odot$ in the direction of rotation and $W_\odot$ towards the North Galactic Pole. Our adopted values for these parameters are given in Table 2, the caption of which contains the relevant references. Given this information, we can determine actual GRVs using the relation

$$\mathrm{GRV_{obs}} = \mathrm{HRV_{obs}} + \boldsymbol{v}_\odot \cdot \widehat{\boldsymbol{d}}_{\mathrm{MW}} \tag{5}$$

We use $\widehat{\boldsymbol{d}}_{\mathrm{MW}}$ for the direction from the MW towards a target galaxy. The term $\left(\boldsymbol{v}_\odot \cdot \widehat{\boldsymbol{d}}_{\mathrm{MW}}\right)$ represents a correction for Solar motion within the MW. Because $v_{c,\odot}$ is a model parameter, this correction is slightly model-dependent.

At large distances from the MW, it and M31 may be considered as a single-point mass M. However, even in this case, the MW mass fraction $q_{MW} \equiv \frac{M_{\mathrm{MW}}}{M}$ has a substantial effect on GRVs. This arises because smaller values of q_{MW} imply that the MW is moving faster with respect to the LG barycentre. As a result, even a spherically symmetric model of the LG can be used to place meaningful constraints on q_{MW}, as was recently done by Peñarrubia et al. (2014).

We have implicitly assumed that the motion of the Sun with respect to the disc of the MW is the same as its motion with respect to the MW barycentre, the important quantity for our timing argument analysis. This assumption may be invalidated if the MW has massive satellite galaxies. In fact, this does seem to be the case, especially when considering the LMC (Peñarrubia et al. 2016).

In our models, the LMC is not treated as another particle but as part of the MW. Thus, its simulated mass includes that of the LMC. Effectively, our model uses one particle to represent the MW system ($\equiv$ MW + satellite). This assumes that the LMC is bound to the MW, whose disc must then be moving with respect to the barycentre of the MW system due to the recoil induced by the LMC. Because it is very nearby compared to other LG galaxies of interest, we neglect the fact that observations made from near the Sun are no longer made at the barycentre of the MW system.[3]

Although such position effects are negligible, it is important to consider the motion of the LMC. Thus, we determined $\boldsymbol{v}_{\mathrm{LMC}}$, its space velocity with respect to the MW. This requires knowledge of its HRV (McConnachie 2012) and its proper motion (Kallivayalil et al. 2013) multiplied by its distance (Pietrzyński et al. 2013). This information was used to obtain a revised estimate for the motion of the Sun with respect to the MW system

$$\boldsymbol{v}_\odot \rightarrow \boldsymbol{v}_\odot - q_{\mathrm{LMC}} \boldsymbol{v}_{\mathrm{LMC}} \tag{6}$$

$$q_{\mathrm{LMC}} \equiv \frac{M_{\mathrm{LMC}}}{M_{\mathrm{MW}}} \quad \text{(MW includes LMC)} \tag{7}$$

Although M31 may have massive satellites too, we do not consider them because they do not affect our analysis to the same extent. A massive satellite of M31 can create a mismatch between the present GRV of the M31 disc and that of the M31 system. Our model would account for the mass of the M31 satellite as part of the M31 mass itself. The main effect on galaxies other than M31 would be a shift in the barycentre of the M31 system by a few kpc. This should have only a very small effect on the dynamics of other galaxies due to their large distances from M31 (top panel of Fig. 1). Consequently, only the GRV of M31 could be noticeably affected in this scenario. As M31 is only one of our 34 target galaxies, our analysis should not change much overall. This is not true if the MW

[3] We take the Sun to be at the barycentre of the MW system. However, our 3D model (Section 3) accounts for the LMC directly and treats the Sun as being 8 kpc from the centre of the MW disc.

had a massive satellite as that would affect our velocity relative to everything else.

2.2 Statistical analysis

We used our axisymmetric model to predict GRVs of target galaxies where it was possible to obtain a unique prediction. This is not always the case, as is clear from Fig. 1. In the region between the MW and M31, intersecting trajectories are apparent. This means that there is more than one possible velocity at the same position, even in the same model. Thus, we do not have any targets within this region. Based on this consideration, we made some adjustments to the catalogue of galaxies used by Peñarrubia et al. (2014) for our analysis in Banik & Zhao (2016). In particular, we excluded Andromeda XVIII and treated HIZSS 3A & B as one bound object, assuming a mass ratio of 13:1 (Begum et al. 2005).

To handle distance uncertainties, we recalculated GRV predictions with each target moved along the line of sight to the 1σ upper limit of its observed distance $d_{\rm MW}$ (using the 1σ lower limit instead yielded almost identical results).

$$\sigma_{\rm pos} \equiv \left| GRV_{\rm model}\left(d_{\rm MW}+\sigma_{d_{\rm MW}}\right) - GRV_{\rm model}\left(d_{\rm MW}\right)\right| \tag{8}$$

Here, $\sigma_{d_{MW}}$ is the uncertainty in the distance to a target galaxy whose most likely distance is d_{MW}. We can now combine HRV measurement uncertainties σ_{v_h} with those on GRV predictions caused by distance uncertainties. Thus, the contribution to the χ^2 statistic from any galaxy i is

$${\chi_i}^2 \equiv \left(\frac{GRV_{\rm model}-GRV_{\rm obs}}{\sigma}\right)^2 \quad \text{where} \tag{9}$$

$$\sigma = \sqrt{{\sigma_{\rm pos}}^2+{\sigma_{v_h}}^2} \tag{10}$$

Uncertainty in the distance to M31 has other subtle effects on our analysis. The gravitational field in the LG would be altered if M31 was at a different distance than the assumed 783 kpc. However, we neglect such effects because, towards the edge of the LG, the only relevant factors are the masses of the MW and M31. In any case, the rather small uncertainty in its distance of 25 kpc (McConnachie 2012) is unlikely to affect our results much because the closest target galaxy to M31 is still $\gtrsim$ 200 kpc from it (top panel of Fig. 1). None the less, this effect is included directly in our 3D model (Section 3).

We used a grid method to explore the parameter space spanned by the total MW and M31 mass, the fraction $q_{\rm MW}$ of this in the MW, the LSR speed $v_{c,\odot}$ and the LMC mass. We use uniform priors on model parameters except $v_{c,\odot}$, for which we assume a Gaussian prior of 239 ± 5 km s^{-1} (McMillan 2011) and add a corresponding contribution to the total χ^2. Thus,

$$\chi^2 = \left(\frac{v_{c,\odot}-v_{c,\odot,\rm nominal}}{\sigma_{v_{c,\odot}}}\right)^2 + \sum_{\substack{\text{Target}\\ \text{galaxies}}} {\chi_i}^2 \tag{11}$$

Because $\sigma_{\rm pos}$ varies slightly with the model parameters, our error budgets become model-dependent. Thus, the best-fitting model is not just that which minimizes χ^2. We quantify the relative probabilities of different models using

$$P(\text{Model}\,|\,\text{Observations}) \propto \mathrm{e}^{-\frac{\chi^2}{2}} \prod_i \frac{1}{\sigma_i} \tag{12}$$

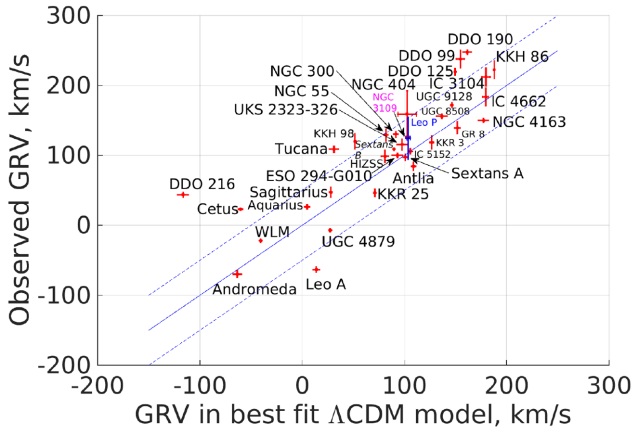


Figure 2. Comparison between predicted and observed GRVs of indicated galaxies in our best-fitting axisymmetric model including Centaurus A and the LMC. The adopted model parameters are given in Table 2. The line of equality is shown in solid blue. Two parallel lines (dashed blue) offset by 50 km s^{-1} are also shown. Assuming our model is accurate to $\sim$25 km s^{-1}, it is unlikely in the context of ΛCDM to find many galaxies far outside this range. Generally, a larger GRV indicates a larger distance (for reference, Aquarius is $\sim$1 Mpc from the LG barycentre).

We will focus on how observed GRVs deviate from those predicted by our best-fitting model. To facilitate the discussion, we define

$$\Delta\text{GRV} \equiv \text{GRV}_{\rm obs} - \text{GRV}_{\rm model} \tag{13}$$

2.3 2D results

In Fig. 2, we compare model-predicted and observed GRVs in the best-fitting model (parameters given in third column of Table 2).[4] We also show results for Cetus and DDO 216. Despite being quite close to M31, they seem to be in a region with a smooth velocity field (Fig. 1). This allows for a well-defined model prediction. However, the absence of M33 from our model may make these predictions less reliable than for other target galaxies. This is especially true with DDO 216, which is closer to M31. We will return to this point later.

There is a tendency for observed GRVs to exceed model predictions (Fig. 2). To gain a better feel for this phenomenon, we constructed a histogram of all the ΔGRVs. Errors due to distance and HRV uncertainties are accounted for by convolving each data point with a Gaussian of the appropriate width (equation 10). To account for an uncertain LSR speed (which is only weakly constrained by our investigation), we also added the 5 km s^{-1} uncertainty on this in quadrature. The results are shown in Fig. 3.

Our model is not a perfect representation of ΛCDM and can only really be expected to get GRV predictions accurate to $\sim$25 km s^{-1} (Aragon-Calvo, Silk & Szalay 2011). Thus, one expects the distribution of ΔGRVs to be broadly consistent with a Gaussian of this width. Indeed, this appears to be the case for galaxies with $\Delta\text{GRV} < 0$ (blue in Fig. 3).

On the other hand, this is not true for galaxies that have $\Delta\text{GRV} > 0$. One can dismiss the bump in the histogram near 160 km s^{-1} due to DDO 216 on the grounds that it may be too

[4] This is similar to fig. 8 of Banik & Zhao (2016), but the LMC is now included at the best-fitting $q_{\rm LMC}$.

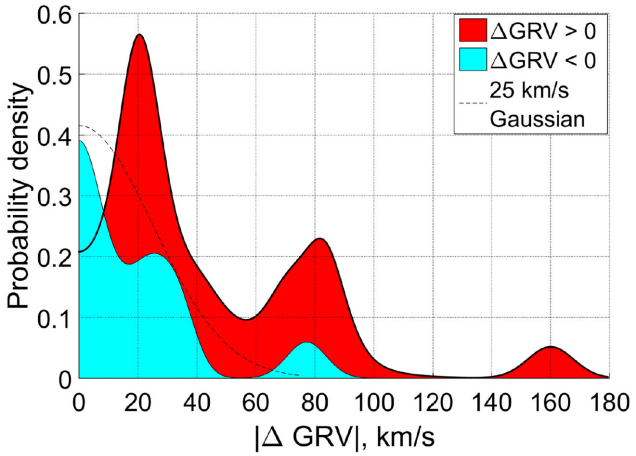


Figure 3. Histogram showing observed – predicted GRVs of our target galaxies using our best-fitting 2D model. The area of each square corresponds to two galaxies. Each data point was convolved with a Gaussian of width $\sigma = \sqrt{{\sigma_{\rm pos}}^2 + {\sigma_{v_h}}^2 + {\sigma_{v_{c,\odot}}}^2}$. This matches the $\Delta{\rm GRV} < 0$ subsample (filled blue) quite well, especially when Leo A is excluded as this removes the blue bump near 75 km s^{-1}. However, it does not match the $\Delta GRV > 0$ subsample (filled red).

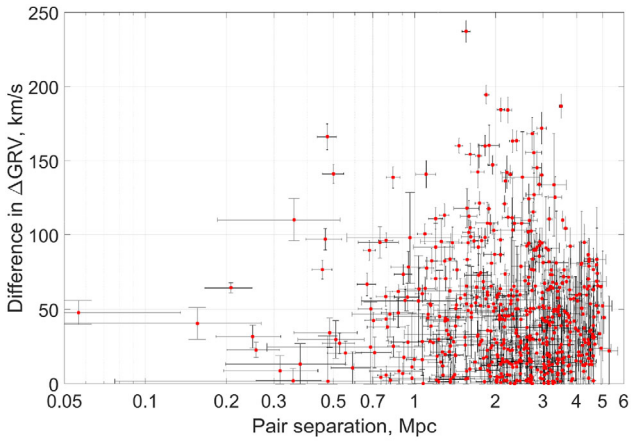


Figure 4. Distances between galaxies in our sample are shown here against the difference between their ΔGRVs. The errors on these correspond to errors on measured distances and RVs of both galaxies being compared. For galaxies near each other, model predictions for their GRVs are similar such that the y-axis essentially just shows the difference between their observed HRVs. Sometimes, these differences are large even for nearby galaxies. These cases appear towards the upper left potion of this figure. The most striking examples are listed in Table 1.

close to M31 and thus the velocity field may be disturbed there. This is not the case in our model (Fig. 1) but one can envisage that it is true in the real world when one considers additional effects, e.g. interactions with massive M31 satellites such as M33. However, it is very difficult to dismiss the bump at 80 km s^{-1} in this way because it corresponds to several galaxies, some of which are quite far from the LG (top panel of Fig. 13). The presence of this feature along with the expected bump near $\Delta{\rm GRV} = 0$ suggests the existence of some additional process responsible for a few galaxies having GRVs much higher than expected in our model. We consider some possible solutions to this high-velocity galaxy problem in Section 5.

At the positions of our target galaxies, we expect the velocity field of the LG to be smooth (Fig. 1). To see if this is the case, we determined the distance between each pair of targets and the difference in their ΔGRVs (Fig. 4).[5]

Some examples are apparent where galaxies are quite near each other but have a very different ΔGRV. In these situations, because model predictions should be very similar, the difference in ΔGRVs must be mostly due to a difference in observed HRVs.[6] More information is given about some of these cases in Table 1.

A few such discrepant pairs are expected given that there are 561 pairs in total. However, the magnitude of the difference between ΔGRVs is rather large in some cases, suggesting that the velocity field in the LG may not be as smooth as in our model. The most convincing examples of this are related to the galaxy NGC 4163, without which the case for a disturbed velocity field is greatly weakened. The galaxies near it (DDO 99, 125 and 190) all seem to have much higher ΔGRVs, suggesting that perhaps the GRV of NGC 4163 is unusually low. In fact, it has the third-lowest ΔGRV of -26.9 ± 7.6 km s^{-1}. This may not seem very low, but we will see later that a more detailed 3D model of the LG predicts a much higher GRV for this galaxy than is observed (Section 3).

Table 1. Differences between the ΔGRV of nearby galaxies are shown here, with our 2D model used to obtain predicted GRVs. The error budgets account for uncertainties in HRV and distance measurements of both targets. Their separation is shown in physical units and as a fraction of the mean of the distances from the MW to each of them. We show only the most extreme examples of galaxies near each other but with a very different ΔGRV (most convincing examples near top). Results for all galaxy pairs are shown in Fig. 4.

Galaxy 1	Galaxy 2	Separation (Mpc)	ratio	Difference in ΔGRV (km s^{-1})
DDO 99	NGC 4163	0.36 ± 0.17	0.13 ± 0.06	110.4 ± 14
DDO 125	NGC 4163	0.47 ± 0.07	0.17 ± 0.03	97.1 ± 7.3
KKR 3	DDO 190	0.74 ± 0.13	0.30 ± 0.05	95.0 ± 10.5
Andromeda	DDO 216	0.47 ± 0.04	0.56 ± 0.05	166.5 ± 8.6
WLM	DDO 216	0.50 ± 0.05	0.54 ± 0.05	141.3 ± 6.5
UGC 8508	DDO 190	0.67 ± 0.05	0.25 ± 0.02	66.9 ± 7.6
WLM	Cetus	0.21 ± 0.04	0.25 ± 0.05	64.5 ± 3.4

A 3D model allows us to test our conclusions more rigorously by directly including many more objects, several of which are quite far from the MW–M31 line. The inclusion of the LMC and M33 can help to make our model more reliable closer to the MW and M31, respectively. This is one reason why we felt comfortable adding Cetus and DDO 216 to our sample. Our model should also be more reliable further from the LG as it now includes more of the most massive objects just outside it. However, we do not take advantage of this by expanding our sample outwards.

3 THE 3D METHOD

3.1 Governing equations

The 3D algorithm we employ is explained in more detail in appendix A of Peebles, Tully & Shaya (2011), which applies the numerical

[5] Errors on mutual separations are over-estimated because we add distance errors in quadrature.

[6] We found it helpful to use ΔGRVs instead of directly observed HRVs as this allows us to include effects such as galaxies further away along the same line of sight being expected to have a higher HRV.

Table 2. The parameters of our best-fitting axisymmetric (2D) and 3D models are given here. $q_{\rm LMC}$ is defined in equation (7). The top section of this table contains the parameters we varied using a grid search in our 2D model (Banik & Zhao 2016) or using a gradient descent method in 3D (Section 3.2). The central section contains the parameters associated with the non-circular motion of the Sun in the MW, which we obtain from Francis & Anderson (2014) for the 2D model and from Schönrich, Binney & Dehnen (2010) for the 3D model. This section also contains two parameters related to M31. In the 2D model, its distance estimate is from McConnachie (2012) while the 3D model uses a prior of 770 ± 40 kpc (Ma et al. 2010). Its rotation curve flatlines at a level $v_{f,\,M31}$, which is fixed in the 2D model but has a prior of 240 ± 10 km s^{-1} in the 3D model (Carignan et al. 2006). This model assumes $v_f = v_{c,\odot}$ for the MW whereas the 2D model fixes the former at 180 km s^{-1} (Kafle et al. 2012) and uses a prior on the latter of 239 ± 5 km s^{-1} (McMillan 2011). We adopt a flat dark energy-dominated cosmology whose parameters are fixed at values given in the bottom section, with the 2D results based on those of Planck Collaboration XIII (2016) while the 3D results are based on Komatsu et al. (2011). Both models start when the cosmic scale-factor $a = 0.1$.

Parameter	Meaning and units	Best-fitting value in 2D	Best-fitting value in 3D
M	LG mass, 10^{12} M$_\odot$	2.756	4.088
$q_{\rm MW}$	$\frac{M_{\rm MW}}{M}$	0.356	0.497
$q_{\rm LMC}$	$\frac{M_{\rm LMC}}{M_{\rm MW}}$	0.157	0.099
$v_{c,\odot}$	LSR speed, km s^{-1}	239	223.0
$v_{f,\,M31}$	v_f of M31, km s^{-1}	225	240.3
d_{M31}	Distance to M31, kpc	783	707
$M_{\rm CenA}$	Cen A mass, 10^{12} M$_\odot$	4	5.883
U$_\odot$	Components of the	14.1	11.1
V$_\odot$	non-circular motion of	14.6	12.2
W$_\odot$	Sun in the MW, km s^{-1}	6.9	7.2
H_0	Hubble constant	67.3	70
$\Omega_{m,\,0}$	Present matter density in the Universe $\div \frac{3{H_0}^2}{8\pi G}$	0.315	0.27

action method to solve the governing equations. A more detailed attempt was later made to use this method to understand the dynamics of LG galaxies (Peebles & Tully 2013). We briefly review some of the key aspects of how the model works.

We adapted a previous dynamical model of the LG based on the objects included in Peebles et al. (2011, Table 1). This brightness-based catalogue is similar to the massive galaxies used in our analysis (Table 3). However, it misses the vast majority of the galaxies analysed in Banik & Zhao (2016), which is a major shortcoming because LG dwarfs ∼1–3 Mpc away turned out to be crucial to its conclusions. Thus, if not already present in our 3D model, we added the galaxies analysed in that work to it as test particles satisfying the equation of motion

$$\ddot{\boldsymbol{r}} = {H_0}^2\Omega_{\Lambda,0}\boldsymbol{r} - \sum_{\substack{\rm j=Distant\\ \rm massive\\ \rm particles}} \frac{GM_j\left(\boldsymbol{r}-\boldsymbol{r}_j\right)}{|\boldsymbol{r}-\boldsymbol{r}_j|^3} - \sum_{\substack{\rm j=Nearby\\ \rm massive\\ \rm particles}} \frac{GM_j\left(\boldsymbol{r}-\boldsymbol{r}_j\right)\left({r_c}^2+{r_{S,j}}^2\right)}{\left(|\boldsymbol{r}-\boldsymbol{r}_j|^2+{r_c}^2\right){r_{S,j}}^3} \quad (14)$$

Table 3. Data on the massive galaxies in our 3D model using a similar catalogue to Peebles et al. (2011, Table 1). Distances and masses are allowed to vary to best match observations, though their prior distributions are not uniform (see text). The masses derived in our model correspond to the total halo mass of each system, some of which is located beyond its virial radius (Peñarrubia & Fattahi 2016). The top section of this table contains galaxies that are also directly included as massive extended objects in our 2D model (Section 2). The remaining galaxies are sorted in descending order of simulated mass. For clarity, we abbreviated the names of galaxies from the New General Catalogue (NGC) and Index Catalogue (IC).

Galaxy	Distance, Mpc	HRV, km s^{-1}	Mass, 10^{12} M$_\odot$
MW	0.008	− 11.10	1.8302
Andromeda (Messier 31)	0.707	− 309.18	2.0567
Centaurus A	3.736	504.52	5.8831
Messier 101	7.391	439.62	9.3108
Messier 94	4.366	324.31	8.8144
Sculptor	4.095	246.97	6.9296
NGC 6946	5.859	107.38	4.6142
Messier 81	3.625	73.48	4.0625
Maffei	3.988	− 28.75	3.4924
IC 342	3.350	− 12.98	1.2994
Triangulum (Messier 33)	0.948	− 192.72	0.2214
Large Magellanic Cloud	0.065	235.97	0.2007
NGC 55	2.035	163.16	0.1323
NGC 300	1.963	158.70	0.1073
IC 10	0.781	− 338.02	0.0437
NGC 185	0.706	− 213.37	0.0129
IC 5152	1.878	138.56	0.0094
NGC 147	0.679	− 201.04	0.0064
NGC 6822	0.510	− 69.93	0.0059

When determining the force between any pair of massive galaxies, the value of r_s used corresponds to the galaxy with the larger r_s. The massive galaxies in this analysis are given in Table 3. The distances and HRVs shown are best-fitting values obtained based on trying to match all observational constraints within their uncertainties (Section 3.2).

The gravitational field near massive particles is handled slightly differently than in our 2D model. For any given test particle A, an explicit distinction is now drawn between massive particles whose r_s is below the distance to A and masses for which this is not the case, forces from which are handled using a pure inverse square law. Forces from nearby masses at first rise linearly with separation before falling as $F \propto \frac{1}{r}$, recovering the observed flat rotation curves of galaxies. The transition occurs around $r_c = 10$ kpc.

For the MW and M31, we define r_s in the same way as previously, though we now add the assumption that the LSR speed is the same as v_f for the MW. Its value is allowed to float, with a prior assumption of 240 ± 10 km s^{-1}. We use the same value for M31. For other massive galaxies, we assume $r_s = 100$ kpc to avoid an adjustment each time their masses are altered.

Some differences with equation (4) are apparent. The part of the cosmological acceleration ä caused by dark energy is handled in the same way but the part caused by matter is not. Previously, we treated the Universe as homogeneous except for a few massive particles. This meant that, without these particles, we would need to recover the cosmic expansion $\boldsymbol{r} \propto a$, which is only possible if $\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r}$.

Here, we treat the Universe as empty except for the massive particles that we explicitly include. Because the Universe is homogeneous on large scales, an accurate understanding of all the mass interior to a sufficiently distant test particle also leads to its

separation from us changing with time as $\boldsymbol{r} \propto a$. To see if this applies to our model, we determined how much mass was in our simulation out to the distance of M101, the most distant galaxy in our sample. The result of 4.9×10^{13} $M_\odot$ corresponds to a sphere of radius 7.01 Mpc filled with matter at a density equal to the present cosmic mean value. This is similar to the observed distance of M101 (Shappee & Stanek 2011), suggesting that the massive galaxies in our model mimic a smooth distribution on large scales with the correct density. A similar conclusion would be reached if we consider only galaxies within 3 Mpc of the MW.

The equations of motion are solved by adjusting a trial trajectory towards the true one. An incorrect trajectory will have a mismatch between the acceleration along it and that expected due to the gravity of other particles. Thus, at each timestep, the positions of all the particles are adjusted to try and equalize the gravitational field acting on each one with the acceleration $\ddot{\boldsymbol{r}}$ it experiences along its trajectory. This is done assuming both respond linearly to a position adjustment, although only the latter does. Thus, a solution can only be obtained after several iterations, each of which is reliant on a matrix inversion to handle the highly interconnected nature of the problem. Certain shortcuts are taken for test particles because their position has no effect on the forces felt by other particles.

This method of solution is second-order accurate because of the standard finite differencing scheme used to obtain accelerations from a series of discrete positions valid at known times. Due to the large number of particle pairs, an adaptive timestep scheme is impractical. Instead, we adapt the temporal resolution to the problem in a fixed way based on physical considerations. Each timestep corresponds to an equal increment in the cosmic scale-factor a. We use 500 steps between when $a = 0.1$ and the present time ($a \equiv 1$).

To check if we have adequate resolution, the problem is solved using forward integration, instead with 5000 timesteps equally spaced in a. The maximum error in the present position is 0.23 kpc while that in the velocity is 0.84 km s^{-1}. Both errors are very small, suggesting that we have enough resolution. Some other checks are also done to verify the numerical accuracy of our solution (Peebles et al. 2011, section 2.4).

The Hubble diagram for our best-fitting 3D model is shown in Fig. 5. For comparison, we overlay results from an axisymmetric model using the same model parameters as this 3D model (last column of Table 2). The basic trend of increasing radial velocity (RV) with distance is apparent in both models, even though they are constructed quite differently.

3.2 Statistical analysis

Like our axisymmetric model, our 3D model accurately matches the observed sky positions of target galaxies. However, this is achieved rather differently. Instead of integrating the equations of motion forwards in time and using the Newton–Raphson method to very precisely match the present position, the 3D model integrates backwards in time starting from a position along the line of sight towards a target galaxy. We no longer require agreement between simulated and observed heliocentric distances. Instead, we add a contribution to the total χ^2 of the model if there is a mismatch. Handling distance uncertainties in this way makes error budgets model-independent, reducing the determination of relative model likelihoods to a simple comparison of their χ^2 statistics.

The distance errors σ_d come from observations. For M31, we use a slightly closer and more uncertain estimate (770 ± 40 kpc; Ma et al. 2010). Galaxies outside the LG might be affected by objects

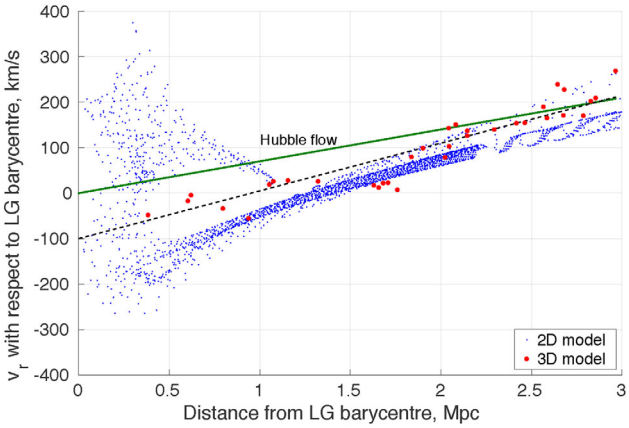


Figure 5. RVs of test particles with respect to the LG barycentre are shown in blue for our 2D model with parameters matched to our best-fitting 3D model (Table 2), results of which are shown as large red dots. The solid green line is the Hubble flow relation for $H_0 = 70$ km s^{-1} Mpc^{-1}, our adopted value. The dashed black line has a gradient 1.5 × larger. Due to the effect of gravity, it provides a better fit to the 3D model within the LG.

beyond the region covered by our analysis. It can also be difficult to determine the mass ratios between galaxies in an extended group and thus the location of its centre of mass. For these reasons, we use a fairly large uncertainty for such distant objects.

$$\frac{\sigma_d}{d_{\rm MW}} = \frac{1}{10} \text{ if } d_{\rm MW} > 3.2 \text{ Mpc} \tag{15}$$

Mismatches between observed and simulated GRVs are handled similarly, based on a tolerance of 20 km s^{-1} rather than the actual HRV measurement uncertainty. This is because we do not expect our model to be much more accurate as a representation of ΛCDM considering the level of scatter about the Hubble flow in more detailed simulations (Aragon-Calvo et al. 2011). As σ_{v_h} is always much smaller than this, the effect of raising it to 20 km s^{-1} is similar to adding an extra 20 km s^{-1} dispersion term to equation (10). This prevents the model placing undue statistical weight on a galaxy with very precise observations, given that the model itself also has uncertainties.

We made use of proper motion data for M31, M33, the LMC, IC10 and Leo I. This was done by adding a penalty to χ^2 when simulated and observed values disagree, with observational error estimates taken at face value.

Unlike in our 2D model, equation (1) is no longer strictly enforced at the start of our simulations because this is difficult to achieve when integrating backwards. Instead, we penalize models that fail to enforce it.

$$\Delta\chi^2 = \frac{|\overbrace{\boldsymbol{v}_i - H_i \boldsymbol{r}_i}^{\boldsymbol{v}_{\rm pec}(t=t_i)}|^2}{{\sigma_v}^2} \tag{16}$$

We assume that the typical peculiar velocity $\boldsymbol{v}_{\rm pec}$ when $a = 0.1$ was $\sigma_v = 50$ km s^{-1} based on present-day deviations from the Hubble flow (Fig. 5). This is a 1D measure that underestimates typical values of $v_{\rm pec}$ today. However, the nearly homogeneous state of the Universe at recombination (Planck Collaboration XIII 2016) implies that $v_{\rm pec}$ is typically larger at the present time than when our simulations started. The distribution of $v_{\rm pec}$ at that time is shown in Fig. 6 for our best-fitting 3D model. All 50 galaxies in this model are represented, some of which are outside the LG.

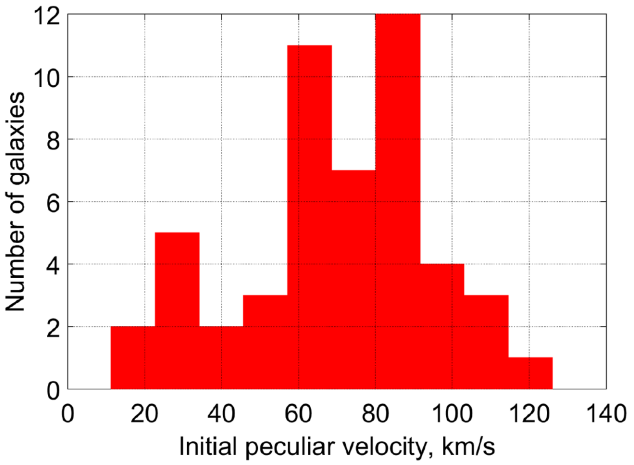


Figure 6. Histogram showing deviations from the Hubble flow (equation 1) at the start of our best-fitting 3D simulation. We allow a tolerance of 50 km s^{-1} (equation 16). The mean peculiar velocity then was 69 km s^{-1}, with a 1σ confidence interval of 38–91 km s^{-1}.

We do not fix the masses of any of our simulated galaxies that are treated as massive (Table 3). However, the prior we use prefers a particular value based on assuming a mass-to-light ratio of 50 times the Solar value in the near-infrared K-band (Tully et al. 2013). Observational estimates of the luminosity in this band L_K are based on a particular distance to each target. If, e.g. the distance in the model is less, then the model implies that the target is likely closer to us and thus intrinsically fainter for the same apparent magnitude. This makes it likely to be less massive. Accounting for this, we define the preferred mass estimate

$$M_c \equiv 50 L_K \left(\frac{M}{L_K}\right)_\odot \left(\frac{d_{\rm model}}{d_{\rm obs}}\right)^2 \tag{17}$$

Using a different mass M incurs an extra χ^2 cost of

$$\Delta\chi^2 = \left[\frac{\mathrm{Ln}\left(\frac{M}{M_c}\right)}{\mathrm{Ln}\ 1.5}\right]^2 \tag{18}$$

For the MW and M31, a slightly different procedure is used. There is no a priori preference towards any particular mass for either galaxy, but a particular *ratio* between their masses is preferred. This is the ratio of their values of M_c.

$$\Delta\chi^2 = \left(\frac{\mathrm{Ln}\ \frac{M_{\rm MW}}{M_{M31}} - \mathrm{Ln}\ \frac{M_{\rm MW,c}}{M_{M31,c}}}{\mathrm{Ln}\ 1.25}\right)^2 \tag{19}$$

Our model now has too many parameters to permit a grid search through them. Thus, we only present results from our best-fitting 3D model. This is obtained by minimizing χ^2 using a downhill-seeking walk through parameter space (Peebles et al. 2011, section 2.2). Each parameter A is varied by a small amount δA in an attempt to reduce χ^2. If this does not happen, then the algorithm restores the previous solution and sets

$$\delta A \rightarrow -\frac{1}{2}\delta A \quad \left(\chi^2 \text{ increased}\right) \tag{20}$$

As well as reversing the sign of δA, it is important to reduce its magnitude because the increase in χ^2 is often caused by overshooting its minimum with respect to A.

When we are fortunate in that a parameter adjustment reduces χ^2, we accelerate the convergence by setting

$$\delta A \rightarrow \frac{5}{4}\delta A \quad \left(\chi^2 \text{ decreased}\right) \tag{21}$$

To avoid the parameter adjustments being too large or too small, a cap and floor are imposed on $|\delta A|$ such that

$$10^{-5} < \left|\frac{\delta A}{A}\right| < 10^{-1} \tag{22}$$

4 3D RESULTS

We started with a solution to the equations of motion that best matched available constraints on the massive galaxies in our analysis (Table 3). We then added test particles one at a time, giving the algorithm an opportunity to adjust the masses and trajectories of the massive objects in order to better accommodate observational constraints from the corresponding LG dwarf. Once we had included our complete sample, we varied all the parameters one at a time to see if we could achieve any further reduction in χ^2. After repeating this a few times, it became clear that the preferred solution was not changing. Some of the most important parameters associated with this model are shown in Table 2.

The optimal value of $q_{\rm MW}$ is 0.50, slightly higher than the 0.36 $\pm$ 0.04 preferred by our axisymmetric analysis (error budgets estimated from Banik & Zhao 2016, fig. 18). This is because our 3D model puts M31 at a distance of only 707 kpc (Table 3), much less than the most likely distance of 783 kpc (McConnachie 2012). If the model is to be trusted, then the known apparent magnitude of M31 combined with a closer distance implies that it is intrinsically fainter. This reduces the preferred mass of M31 (equation 17). If we scale up the mass of M31 by $\left(\frac{783}{707}\right)^2$, then $q_{\rm MW}$ would fall to 0.45. Because we include the LMC as part of the MW when determining $q_{\rm MW}$, it seems reasonable to treat M33 as part of M31 rather than as a separate object. Doing so reduces $q_{\rm MW}$ a further $\sim$0.03. This makes it consistent with our axisymmetric analysis, assuming that both yield similarly uncertain estimates of q_1.

We consider several galaxies close enough to M31 for the flatline level of its rotation curve to make some difference. This is especially true with NGC 147 and NGC 185. To a lesser extent, it is also the case for IC 10. This is interesting in light of its measured proper motion (Brunthaler et al. 2007). Thus, our model may be able to constrain $v_{f,M31}$. It prefers a value of 240.29 km s^{-1}, very close to the 240 km s^{-1} at which the prior distribution peaks (Table 2). We are unable to determine the precision with which our model constrains this parameter. We believe that' the suspiciously good agreement (within 0.03σ of the prior) indicates that our analysis is simply unable to obtain meaningful constraints on $v_{f,M31}$, such that its prior is the most important consideration.

Our analysis preferred a low value of $v_{c,\odot}$, so we focused on adjusting only this parameter to better constrain its optimal value. This did not alter our results. Our 3D analysis alone must prefer even lower values than the best fit, which also considers our prior of 240 $\pm$ 10 km s^{-1}. Given that analyses such as these typically constrain $v_{c,\odot}$ to within no better than $\sim$15 km s^{-1} (e.g. Peñarrubia et al. 2014; Banik & Zhao 2016), it appears that there is some tension between the 223 km s^{-1} preferred by our analysis and the independent[7] estimate of 239 $\pm$ 5 km s^{-1} (McMillan 2011). Interestingly, a more recent estimate preferred a lower value of

[7] not based on the timing argument.

232.8 ± 3.0 km s^{-1} (McMillan 2017). However, the reduced uncertainty leads to the same conclusion.

Our model also contains some galaxies quite close to the MW, making it important to have an accurate force law within its virial radius. We may be failing to achieve this by assuming $v_{c,\odot} = v_f$, the speed at which the rotation curve of the MW flatlines. It is possible that the best-fitting LSR speed obtained by our algorithm has been dragged down because $v_f \ll v_{c,\odot}$, as suggested by Kafle et al. (2012). We hope to relax the assumption that $v_{c,\odot} = v_f$ in a future investigation.

Our axisymmetric analysis had almost no preference for an LSR speed different to the 239 km s^{-1} peak of its prior distribution (Banik & Zhao 2016, table 2). The lower value of $v_{c,\odot}$ preferred by our 3D analysis affects M, the inferred total mass of the MW and M31. This is because M31 is almost directly ahead of the Sun in its orbit around the MW. Thus, a lower LSR speed implies that more of the observed blueshift of M31 must be due to it moving towards the MW, requiring a higher combined mass for these galaxies. We estimate that a 16 km s^{-1} reduction in $v_{c,\odot}$ increases M by $\sim 0.8 \times 10^{12}\,\mathrm{M}_\odot$ (Banik & Zhao 2016, fig. 7). A higher M is likely also required to counteract the stronger effect of tides raised by Cen A due to its higher inferred mass. Its location very close to the MW–M31 line and relative proximity make it an important consideration. Additionally, we expect a similar effect due to the different assumptions concerning the background density of matter in the LG. Our 2D analysis assumed that the LG was filled with matter at the cosmic mean density. In our 3D model, it is treated as empty apart from a few discrete objects like the MW and M31. Our previous results suggest that this should increase the best-fitting value of M by $\sim 10^{12}\,\mathrm{M}_\odot$ (Banik & Zhao 2016, section 4.1). For these reasons, it is not too surprising that our 3D analysis prefers a higher M, even though this is counteracted slightly by the lower preferred distance to M31.

We conducted versions of our axisymmetric analysis without Cen A and with it included at a mass of $4 \times 10^{12}\,\mathrm{M}_\odot$. The latter provided a much better fit to observations (Banik & Zhao 2016, fig. 14). Our 3D algorithm starts off with Cen A having $M_c = 1.06 \times 10^{13}\,\mathrm{M}_\odot$ but the analysis prefers a value of just over half this (not problematic given the fairly broad mass priors – see equation 18). This suggests the possibility of constraining the masses of galaxies just outside the LG based on their tidal effect within it. Such an analysis is likely to face degeneracies between masses of different galaxies along a similar line of sight, but it might still be worthwhile.

A comparison between our best-fitting 2D and 3D models is complicated somewhat by the latter having many more degrees of freedom. In particular, it is not required to match the observed distances of LG galaxies. This allows it to place a galaxy further away than observed, increasing its predicted GRV and better explaining a very high observed GRV. We handle this by applying a correction to model-predicted GRVs if they correspond to a simulated galaxy at a different distance than the real one it is supposed to represent. Thus, we set

$$\mathrm{GRV}_{\mathrm{model}} \rightarrow \mathrm{GRV}_{\mathrm{model}} + (d_{\mathrm{obs}} - d_{\mathrm{model}})\,\alpha H_0 \tag{23}$$

$$\alpha \equiv \frac{1}{H_0}\frac{\mathrm{d}v_r}{\mathrm{d}r} \tag{24}$$

We use $\alpha = 1.5$ because this seems to provide a reasonable description of how RVs v_r depend on distances within the LG (Fig. 5). At long range, we would get $\alpha = 1$. Within the LG, gravity from the MW and M31 is important. Thus, an object further from them has been decelerated less by their gravity. This means that its RV will be higher by a greater amount than in a homogeneously expanding Universe. Neglecting projection effects (which become small a few Mpc from the LG), we see that α should slightly exceed 1.

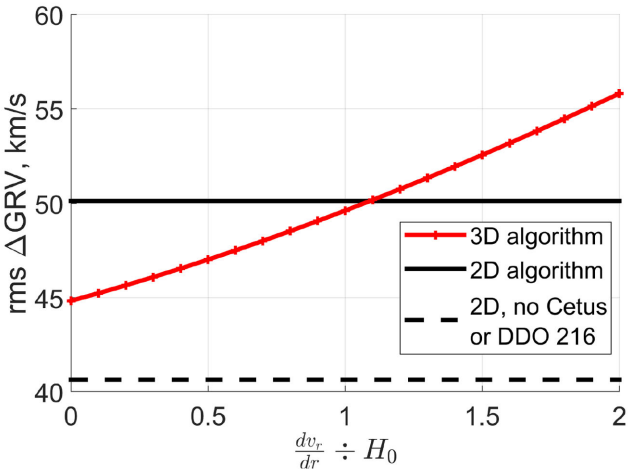


Figure 7. The root mean square value of ΔGRV is shown here for the best-fitting 2D (black) and 3D (red) models. α governs the way we adjust 3D model predictions to put them on an equal footing with our 2D model (equation 23). The adjustment is unnecessary for the latter (see text). This model is likely unreliable close to M31 as it lacks M33, making its predictions for Cetus and DDO 216 unreliable. Results including these galaxies (solid black) and without them (dashed black) are shown. In the 3D model, removing them increases the results by ~ 0.7 km s^{-1}, leaving them almost unchanged. This model treats the LG as empty apart from a few point masses. Using a similar assumption in our 2D models would reduce the rms value of ΔGRV by ~ 6 km s^{-1} (not shown).

In Banik & Zhao (2016), we added an extra dispersion term to equation (10) and marginalized over other variables to obtain a probability distribution for it. The most likely value we obtained (using the optimal LMC mass) was 40.43 km s^{-1}. Using the same target galaxies, the rms dispersion in ΔGRV with respect to the best-fitting 2D model is 40.65 km s^{-1}, almost exactly the same. This suggests that the two statistics are very similar, even though the former uses integration over model parameter space while the latter is based on just one model. Thus, the rms ΔGRV likely provides a very good guide to the results of a more thorough statistical analysis attempting to pin down how inaccurate our model is as a representation of the data.

After obtaining corrected GRV predictions using equation (23), we subtract them from observed GRVs (equation 5) to obtain a list of ΔGRVs. We determine the rms of these ΔGRVs for a range of plausible assumptions regarding α. The precise value does not much affect our overall conclusions (Fig. 7).

For comparison, we also show the result of the same calculation for our best-fitting 2D model using the same target galaxies. This model requires an extremely precise match between their simulated and observed distances, making the result independent of α. Although it was technically difficult for us to operate the 3D model in this way, we can gain a conservative lower bound on the rms value of ΔGRV if we did this by considering the $\alpha = 0$ case in Fig. 7. This corresponds to taking the GRV predictions of the 3D algorithm at face value, even though it has some flexibility with distances. Removing this flexibility can only worsen the agreement between predicted and observed GRVs.

With α irrelevant for our 2D model, the main uncertainty becomes whether Cetus and DDO 216 are included in the analysis as they are very discrepant with this model. We suggest that they should not be included as they are quite close to M31 (Fig. 1). If one also excludes them from the 3D analysis, then the rms value of ΔGRV for it is hardly affected (it rises ∼0.7 km s^{-1}), thus greatly exceeding the 2D result for the same sample. Even if these galaxies are included, any value of $\alpha > 1.1$ implies that the rms ΔGRV is larger in our 3D analysis.

The results shown in Fig. 7 for the 2D model correspond to an LG filled with matter at the cosmic mean density. To see how much this assumption might affect our results, we previously repeated some of our 2D calculations assuming an empty LG apart from the MW and M31 (Banik & Zhao 2016, section 4.1). This naturally raises predicted GRVs towards the outskirts of the LG, thereby improving the agreement with observations and reducing the extra dispersion by ∼6 km s^{-1}. This reinforces our conclusion that the 3D model does not yield a better match to observations than our previous axisymmetric investigation. In fact, the agreement is slightly worse for the most plausible model assumptions.

Our previous work suggests that we can obtain an error estimate for Fig. 7 using the usual rule for the uncertainty in the rms of N-independent random variables. In this case, the fractional uncertainty when $N \gg 1$ is $\frac{1}{\sqrt{2N}}$, where the number of galaxies is $N = 34$. Thus, we expect an uncertainty of ∼6 km s^{-1}, making at least the 3D results inconsistent with the 30 km s^{-1} scatter about the Hubble flow found by Aragon-Calvo et al. (2011).

In order to estimate uncertainties more rigorously, we run another axisymmetric simulation with parameters chosen to match those in our best-fitting 3D model. This is only possible for some parameters (shown in Table 2) as several relate to particles unique to the 3D model and to motion in 3D.

Using this model, we obtain estimates of how much uncertainty there is on the predicted GRV of each target because of its uncertain position along the line of sight (equation 8). To obtain the uncertainty on its ΔGRV, we also need to add in quadrature the uncertainty on its observed RV σ_{v_h} (equation 10).[8] We expect this to capture the major observational sources of error.

After determining ΔGRV and its uncertainty for each target with respect to our 3D model, we can readily see if any galaxies have an unusual redshift for their position. Neglecting a couple of galaxies with large distance uncertainties,[9] we show the five most extreme cases of anomalously high and low GRVs in Table 4. It is apparent that several galaxies have observed GRVs substantially different from that predicted by our best-fitting model. Most of these galaxies have ΔGRV > 0.

We have treated NGC 3109 and Antlia as separate objects. However, they may be gravitationally bound (van den Bergh 1999). There are indications that they have recently interacted, based on observations of both NGC 3109 (Barnes & de Blok 2001) and Antlia (Penny et al. 2012). This is more likely if Antlia is a satellite of NGC 3109. The 41 ± 1 km s^{-1} difference in their GRVs and their 1.°19 sky separation (corresponding to $\gtrsim$28 kpc) are likely consistent with this scenario if their heliocentric distances are similar. The distance to Antlia was found to be 1.31 ± 0.03 Mpc by a study focusing exclusively on this galaxy (Pimbblet & Couch 2012), similar to the 1.286 ± 0.015 Mpc measured previously by Dalcanton et al. (2009). The range of published distances to NGC 3109 is wider than their formal uncertainties, but the most accurate one (based on Cepheid variables) is 1.300 ± 0.012 Mpc (Soszyński et al. 2006). Thus, these two galaxies are probably not much more than 40 kpc apart and may well be bound.

Table 4. ΔGRVs with respect to our 3D model for the LG galaxies most discrepant with it (excluding NGC 404 and Leo P due to large distance uncertainties, see text). Errors are estimated using equation (10). The LG barycentre is put almost exactly at the MW–M31 mid-point (Table 2). Errors in the distance from there are obtained from those on heliocentric distances in the usual way.

Galaxy	ΔGRV (km s^{-1})	Distance from LG barycentre (Mpc)
HIZSS 3	123.2 ± 10.6	1.76 ± 0.11
NGC 3109	110.7 ± 7.3	1.63 ± 0.05
Sextans A	95.1 ± 7.2	1.66 ± 0.02
Sextans B	75.4 ± 5.4	1.71 ± 0.05
Antlia	61.6 ± 8.3	1.68 ± 0.06
UGC 4879	−31.1 ± 5.5	1.32 ± 0.02
KKR 3	−33.6 ± 10.9	2.30 ± 0.12
GR 8	−40.0 ± 10.5	2.42 ± 0.12
NGC 55	−42.0 ± 10.4	2.08 ± 0.11
NGC 4163	−130.6 ± 7.7	2.96 ± 0.04

Even if it turns out that Antlia should not be treated as an independent object, our overall conclusions should not be much affected because its ΔGRV is close to the typical ∼50 km s^{-1} (Fig. 7). As a result, the removal of Antlia[10] from our sample only reduces the rms value of ΔGRV by 0.30 km s^{-1}. In this case, the appearance of Table 4 would also remain similar, with Tucana taking the place of Antlia. Tucana has a ΔGRV of 60.3 ± 7.7 km s^{-1} and is located 1.07 ± 0.05 Mpc from the LG barycentre.

We did not put Leo P into Table 4 due to a 32 km s^{-1} uncertainty on its ΔGRV, almost entirely due to a rather uncertain distance of 1.72 ± 0.4 Mpc. This is derived from ground-based observations (McQuinn et al. 2013). However, a more accurate distance measurement has recently been made using the *Hubble Space Telescope* (1.62 ± 0.15 Mpc; McQuinn et al. 2015). Based on how GRV predictions in our axisymmetric model vary with the assumed distance to Leo P, we estimate that this increases its ΔGRV by ∼9 km s^{-1} while reducing the error on it to only ∼13 km s^{-1}.

Considering the large difference between simulated and observed distances to Leo P (ruled out at almost 4σ using the newer distance), we looked more closely into how much we should adjust its GRV prediction to make this correspond to its observed position. Normally, we use equation (23) without worrying too much about the precise value of α as typical distance errors are small. This is not the case here. Assuming that our axisymmetric model (with the same parameters as our 3D model) provides a better guide to how GRV predictions are affected by line-of-sight distances, it appears that α is overestimated slightly for Leo P. A value of only 1.12 is more appropriate, implying that we reduced its GRV prediction too much to account for it being closer in reality than in our model. As a result, its ΔGRV is slightly smaller, with a best guess of 68 ± 13 km s^{-1} using the updated distance and method.

Our closer look into Leo P hardly changes our estimate of its ΔGRV from the 72 ± 32 km s^{-1} assumed in the rest of this work. However, the error budget is more than halved, making it as discrepant with our 3D model as Antlia and Tucana. Importantly, the

[8] The actual observational uncertainty is used, not 20 km s^{-1}.

[9] Leo P and NGC 404 have ΔGRVs of 72 ± 32 and 70 ± 38 km s^{-1}, respectively.

[10] almost 5 mag fainter than NGC 3109 (McConnachie 2012, table 3).

small change to the result for Leo P and the almost negligible effect of removing Antlia from our sample both lend confidence that our results should not change too much with future improvements to the data and model. This is especially true when one considers that the data for Leo P are particularly inaccurate if using its old distance estimate (Fig. 13). Almost all other galaxies have substantially smaller observational uncertainties.

5 DISCUSSION

Realizing the difficulty faced by our axisymmetric model in explaining observations of the LG, we used a 3D model with many times more free parameters (Table 3). The model was also constrained using more observations, but a lot of these were given quite large uncertainties. For example, we gave all simulated galaxies an HRV error budget of 20 km s^{-1} and assumed a 10 per cent distance uncertainty for galaxies outside the LG (equation 15). We also relaxed the requirement for galaxies to be following the Hubble flow (equation 1) at the start of our simulations. It is clear that our best-fitting 3D model has taken full advantage of this liberty (Fig. 6).

As well as relaxing the constraints on our model in these ways, we used a slightly higher value for the Hubble constant to take advantage of a previous simulation and obtain convergent results faster. We also treated the effect of cosmological acceleration differently (equation 14). In our axisymmetric model, these changes generally increase the predicted GRVs of galaxies near the edge of the LG for fairly intuitive reasons, thus improving the fit to observations (Banik & Zhao 2016, sections 4.1 and 4.2).

Despite these alterations, our best-fitting 3D model provides a similarly poor match to observations as our best-fitting axisymmetric model (Fig. 7). Some dispersion amongst the ΔGRVs is expected due to observational and modelling uncertainties. Using a rigorous grid investigation of the model parameters, we showed previously that there is an unacceptably large typical mismatch between observations and predictions made by our axisymmetric model (Banik & Zhao 2016, section 4). Similar conclusions would be obtained considering the best-fitting model alone (Section 4). We assume that this is also true in our 3D model, justifying our focus on the best-fitting parameter combinations in both models.

The massive galaxies in our models are treated as having a constant mass, implicitly assuming that they formed instantaneously at redshift 9. In reality, galaxies typically gain mass through accretion (Zhao et al. 2009). However, we expect that this should not much affect the present-day velocity of a galaxy required to be at a particular position. This is because the same initial conditions would lead to a different final position, requiring us to adjust the initial conditions so as to counteract this. As a result, the present velocity of a galaxy at a known position is not very sensitive to the forces acting on it at earlier times. Accounting for this 'initial condition drag', the effect on the present velocity of a galaxy due to an impulse at some earlier time is $\propto a^{2.4}$ for an impulse applied when the scale-factor of the Universe was a (Banik & Zhao 2016, fig. 4). To obtain a rough estimate for how much accretion histories might affect our results, we varied the start time of our simulations to correspond to redshift 14 instead of 9 (Banik & Zhao 2016, section 4.6). Because of the initial condition drag effect, the mismatch between model and observations was affected only by $\sim$1 km s^{-1}.

Our results must depend somewhat on the assumed mass profile for our simulated massive galaxies. In each case, the profile is constrained observationally by the flatline level of the observed rotation curve of the corresponding galaxy. With this constraint, whatever specific scheme is used for softening the forces close to massive objects (e.g. equation 14), there must be a transition to an inverse square law beyond a distance of $r_s \sim \frac{GM}{v_f^2}$. This is only $\sim$150 kpc for both the MW and the M31. Thus, there can be little doubt about the force law further than $\sim$1 Mpc from them, the region where the galaxies most discrepant with our model seem to lie (Table 4).

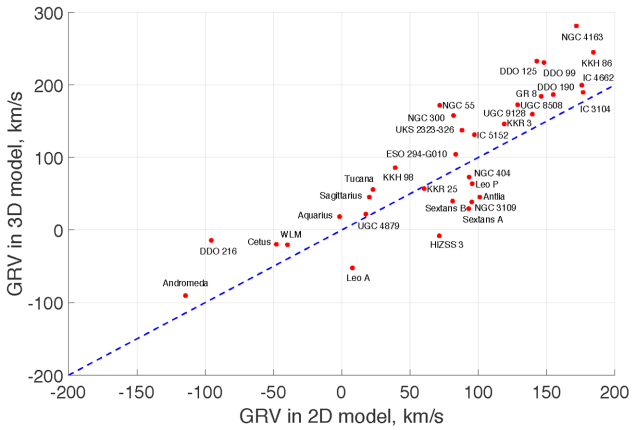


Figure 8. Comparison between the GRVs predicted by our 2D and 3D models with the same model parameters (last column of Table 2). The line of equality is shown in blue. The 2D results may be biased lower towards the edge of the LG because it is assumed to be filled with matter at the same density as the cosmic mean value.

Without a grid investigation of the parameters in our 3D model, one might be concerned whether a qualitatively different type of trajectory for some of the massive galaxies might alter the gravitational field in the LG so as to greatly improve the fit to observations. We consider this unlikely for a timing argument analysis such as ours. The reason is that any plausible solution has the MW and M31 turning around just once, while more distant objects outside the LG have not turned around. It is not feasible for the MW and M31 to have undergone a past close flyby in ΛCDM (or in our model) for any realistic total mass of these galaxies. Thus, the algorithm needs to solve only the continuous problems associated with determining when the turnaround occurred, what the present distance to M31 and its GRV are, etc. It does not need to solve discrete problems like how many times the MW and M31 have turned around. This suggests that gradual adjustments to the model parameters should eventually converge on the best-fitting solution, as assumed in the rest of this work. However, it should be borne in mind that the solution presented here is preliminary, and a better fit to the observations might eventually be obtained using our model.

Although our 2D and 3D models yield broadly similar conclusions, this is not always the case for individual galaxies. To compare the models in a more direct way, we ran our axisymmetric model using the same parameters as our best-fitting 3D model (Table 2), ignoring features unique to the latter. The GRV predictions of our 3D model were referred to the barycentre of the MW and LMC as the latter is included indirectly only in our 2D model (Banik & Zhao 2016, section 4.4).

Predictions from both models are compared in Fig. 8. The models broadly agree, with an average absolute difference of 44.6 km s^{-1}. Interestingly, the 3D results are biased higher by 20.4 km s^{-1}. This may be a consequence of the 2D models treating the LG as filled with matter at a density equal to the cosmic mean value. In the 3D model, it is treated as empty apart from a few point masses. It is unclear how much matter is spread diffusely around the LG,

but clearly 0 is a lower limit. A larger amount almost certainly worsens the agreement with observations (Banik & Zhao 2016, section 4.1). This is because the decelerating effect of the extra matter makes it more difficult to explain the high observed GRVs of several LG galaxies. Thus, the 3D results shown in Fig. 7 may well underestimate the actual extent of the discrepancy between this model and observations.

To estimate how much this assumption might affect the velocity of a target galaxy currently $d = 3$ Mpc away, we determined the time integral of the force on it due to a fixed interior mass M corresponding to the present cosmic mean matter density.

$$M = 4 \times 10^{12}\,\mathrm{M}_\odot \left(\frac{d}{3\mathrm{Mpc}}\right)^3 \tag{25}$$

This is appropriate for a galaxy that is almost following a pure Hubble flow relation at all times, so that its distance in the past is $a(t)d$. Importantly, the effect of forces acting at earlier times needs to be reduced by a factor of $a^{2.4}$ to account for initial condition drag (Banik & Zhao 2016, fig. 4). Bearing this in mind,

$$\Delta v \approx \int_{t_i}^{t_f} \frac{GM}{(ad)^2} \cdot a^{2.4}\,\mathrm{d}t \tag{26}$$

The expansion history of the Universe (equation 3) can be approximated as

$$a\,(t) \approx H_0 t \tag{27}$$

Thus, for simulations starting when $a \equiv a_i$, we get that

$$\Delta v \approx \frac{5}{7}\frac{GM}{d^2 H_0}\left(1 - a_i^{1.4}\right) \tag{28}$$

$$= 19.5\,\mathrm{km\,s^{-1}} \quad \text{for } d = 3\mathrm{Mpc} \tag{29}$$

An effect of this magnitude could go a long way towards explaining why the GRV predictions in our 3D model tend to exceed those in our 2D model. However, there must be other factors as well because a 20 km s^{-1} effect due to the different treatment of mass in the LG is valid only for a galaxy currently 3 Mpc away. In reality, most of our targets are at smaller distances (Fig. 1). As a result, they are less affected by a homogeneous distribution of matter because the Shell Theorem implies that the relevant mass $M \propto d^3$ (equation 25).

One such factor may be that, unlike our 2D model, the 3D model has the freedom to increase predicted GRVs through adjusting the tides raised on the LG by objects outside it (Table 3). Due to the divergence-free nature of the gravitational field far from these perturbers, one also expects RVs to be reduced for LG galaxies in certain directions (equation 30). We suppose that our algorithm tries to ensure that there are fewer target galaxies – especially those with a high ΔGRV – towards these directions.

As well as increased flexibility at long range, our 3D model should be more accurate close to the MW and M31 because it directly includes their most massive satellites. This is probably why it achieves a much better fit to the GRV of DDO 216, which is otherwise too high in our 2D model. In our 3D model, it has a close (73 kpc) encounter with M33 almost exactly 8 Gyr ago.

To gain a feel for the overall pattern of discrepancies between our model and observations, we construct a histogram of the ΔGRVs of our target galaxies using a similar procedure to that used for Fig. 3. This is shown in Fig. 9. Because there is some uncertainty in how to convert HRVs into GRVs due to imperfect knowledge of $v_{c,\odot}$, we add 5 km s^{-1} in quadrature to the other uncertainties. Although the 3D analysis uses a wider (10 km s^{-1}) prior on $v_{c,\odot}$, we choose to stick with the previous value to allow for easier comparison between the histograms.

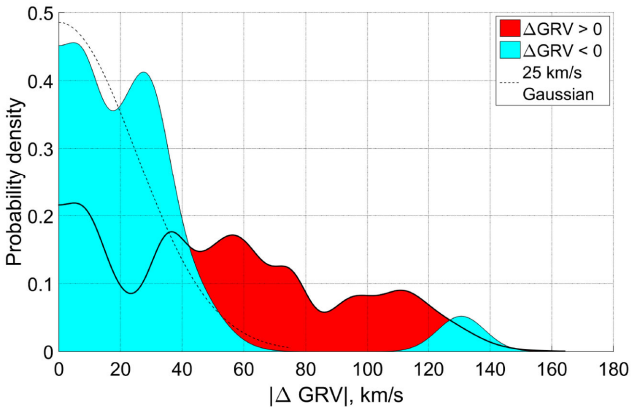


Figure 9. Histogram of ΔGRVs with respect to our 3D model, shown separately according to the sign of ΔGRV. The area of each square corresponds to two galaxies. A similar pattern emerges to Fig. 3, with the blue bump near 130 km s^{-1} caused by NGC 4163. Otherwise, the galaxies with ΔGRV < 0 (filled blue) are well described by a 25 km s^{-1} Gaussian (dashed line). This is not the case for galaxies with ΔGRV > 0 (filled red).

Both models have a tendency for observed GRVs to exceed predicted ones. The pronounced bump near ΔGRV $= 80$ km s^{-1} in the 2D model gives way to a much broader and shallower bump at a slightly lower ΔGRV with respect to our 3D model. Neglecting the bump near 160 km s^{-1} in the 2D results because it corresponds to DDO 216, the distribution of ΔGRVs extends to even higher values in the 3D model.

Some of the galaxies we identify as having a high ΔGRV have previously been identified as such. For example, Pawlowski & McGaugh (2014) identified NGC 3109, Antlia, Sextans A, Sextans B and Tucana as having anomalously high GRVs. With a slightly lower ΔGRV of 60.3 ± 7.7 km s^{-1}, Tucana is still significantly discrepant with our model.

Teyssier, Johnston & Kuhlen (2012) identified these five objects as possible backsplash galaxies, i.e. they have likely passed within the virial radius of the MW or M31. There, they may have passed close to a massive satellite such as M33. Despite allowing for such trajectories, we are still unable to provide a good match to the observed GRVs of several LG dwarf galaxies. We also note that Peebles et al. (2011) found it very difficult to incorporate Sextans A and B into their dynamical model of the LG, suggesting that it could not easily explain their motions. Thus, the high-velocity galaxy problem appears to persist even with a 3D model.

Interestingly, some of these galaxies appear to be correlated in phase space. In particular, the galaxies in the NGC 3109 association (NGC 3109, Antlia, Sextans A, Sextans B and Leo P) seem to lie very close to a line (Bellazzini et al. 2013). Their RVs also closely follow a tight trend of increasing for galaxies further from the LG. We hope to investigate this further in a future publication.

The most extreme outlier with respect to our model is actually a galaxy with negative ΔGRV. This is NGC 4163, a galaxy that appears unusual even in the almost model-independent analysis shown in Table 1. Here, it is apparent that two nearby galaxies have a much higher ΔGRV, suggesting that GRVs of galaxies near NGC 4163 are higher than its GRV. This is also apparent when considering the RVs of other galaxies at similar heliocentric distances in the Canes Venatici I cloud (Makarov, Makarova & Uklein 2013,

table 2). Thus, leaving aside any models, it appears that the HRV of NGC 4163 is ∼100 km s^{-1} less than that of neighbouring galaxies.

One possible explanation is that NGC 4163 was flung towards the LG as a result of a gravitational slingshot interaction in a nearby galaxy group. It is also possible that NGC 4163 had a more recent interaction with a galaxy close to it, as perhaps suggested by its recent starburst (McQuinn et al. 2009). In fact, our 3D model has DDO 99 passing within 50 kpc of it, although this happened in ancient times (when $a = 0.32$). None the less, it is possible that our model has got the timing of this encounter wrong, especially as it does not simulate the effects of any such encounter because both galaxies are treated as test particles. Certainly such a close encounter is rare in our models. If it did happen, then the fact that NGC 4163 is almost a magnitude fainter (McConnachie 2012, table 3) suggests that its dynamics would be affected to a greater extent. Thus, several explanations are possible for its anomalously low GRV.

Interactions amongst galaxies in a neighbouring group can fling a galaxy towards the LG, leading to it having an unusually low GRV. However, it is very difficult to explain a galaxy with anomalously high GRV in this way. This is possible only if the galaxy has crossed the LG and is now heading away from it again. It is not feasible for a galaxy like NGC 3109 to cover such a large distance in the time since the big bang. To see why, consider that it is 1.6 Mpc from the LG, such that the Hubble velocity is ∼110 km s^{-1}. Given an extra RV of 120 km s^{-1}, we see that the galaxy can have covered perhaps twice as much distance as a typical galaxy on the Hubble flow.[11] This means that the anomalous motion of NGC 3109 could not have originated much further from the LG than its present distance, although it could have done so on the opposite side to its current location.

Furthermore, even if this scenario was plausible, it is clear that it would be rarer than a situation where a galaxy is flung towards the LG and we observe it on the way in. Thus, we might see several galaxies with anomalously high GRV, but we would expect to see even more with anomalously low GRV. The opposite seems to be the case (Fig. 9).

Using a 3D model, we can quantify how fast target galaxies move out of the plane they define with the MW and M31. Referring velocities to the barycentre of the MW and LMC, we obtain the results shown in Fig. 10. Part of the reason for motions deviating from axisymmetry with respect to the MW–M31 line is that the direction of this line has rotated slightly because the MW and M31 are not on a purely radial orbit. However, this is only an ∼30 km s^{-1} effect (van der Marel et al. 2012b). Thus, the explanation must lie mostly with the non-axisymmetric gravitational field caused by massive objects far from the MW–M31 line. Massive satellite galaxies such as the LMC and M33 must also play a role.

The substantial non-axisymmetric motions suggested by our analysis are probably required in order to boost the centrifugal force, helping explain the high GRVs of several LG galaxies. Some of this tangential motion may be a relic of peculiar velocities at the start of our simulation (Fig. 6), especially when considering that Hubble drag does not dissipate specific angular momentum (equation 10 of Banik & Zhao 2016). Although the expansion of the Universe naturally reduces tangential speeds, this is not as important an effect as one might think. For example, our best-fitting 3D model has NGC 3109 starting off 593 kpc from the MW and being 1.37 Mpc away currently. This means that its distance has increased <2.5× despite the Universe overall having expanded by a factor of 10.

[11] A more accurate estimate is given in equation (39).

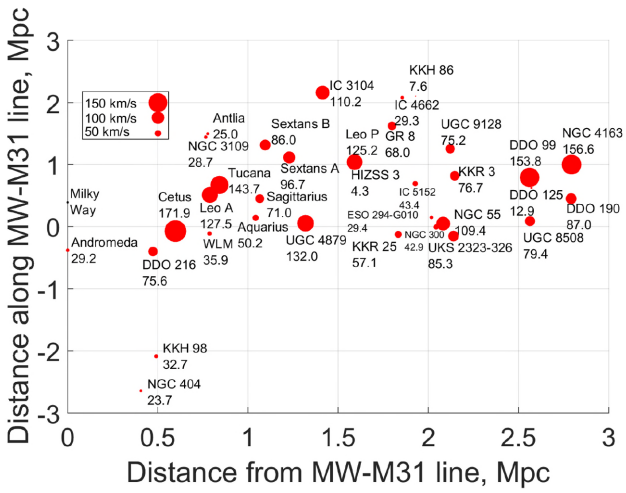


Figure 10. Velocities of our target galaxies out of the plane they define with the MW and M31 are shown here using marker sizes. This information is also given as a number (km s^{-1}) below the corresponding galaxy name. Galaxy positions are shown relative to the MW–M31 line.

In the future, such non-axisymmetric motions should become detectable based on proper motion measurements. A good target for this may be DDO 216 due to its relative proximity. Our model implies that it had a past close encounter with M33, constraining the possible trajectories of these two galaxies and also indirectly forcing some limits on the past position of M31.

The larger expected velocities of Cetus and Tucana out of the plane they define with the MW and M31 should make their proper motions similar in magnitude to that of DDO 216, despite a greater distance. Tucana is one of the nearest galaxies with an unusual GRV in our analysis. But several galaxies with much larger ΔGRV lie just a little further away, in the vicinity of NGC 3109 (Fig. 13). Its proper motion promises to be an extremely interesting constraint on any models of the LG (Pawlowski & McGaugh 2014).

5.1 Effects of the Great Attractor and the Virgo Cluster

A 3D model including several objects beyond the LG allows for a much more rigorous handling of the tides they raise within it. Such effects tend to be larger at greater distances. Thus, it is interesting that the most discrepant objects found by our previous axisymmetric analysis tend to lie towards the edge of the LG (Banik & Zhao 2016, fig. 10). A similar trend is apparent when the LMC is included indirectly (Fig. 2), bearing in mind that a high GRV generally implies a greater distance and that the absence of M33 from the model makes it less reliable for galaxies close to M31 (e.g. Cetus and DDO 216).

In our 3D model, despite handling tides more rigorously, we found a similarly poor match between observed and model-predicted GRVs (Fig. 7). However, the model does not include all the mass concentrations outside the LG that may be relevant to our analysis. An important example is the Great Attractor (GA; Mieske, Hilker & Infante 2005). This is thought to be primarily responsible for the ∼630 km s^{-1} magnitude of $\boldsymbol{v}_{\rm pec,LG}$, the velocity of the LG as a whole with respect to the surface of last scattering (Kogut et al. 1993).

To estimate the effect of the GA on GRVs of objects within the LG, we use the distant tide approximation. Treating LG galaxies as freely falling in the gravitational field of a distant point mass, the

change in the GRV of a target galaxy due to the GA is given by

$$\Delta\mathrm{GRV}_{\mathrm{GA}} = \left(3\cos^2\theta - 1\right)\beta_{\mathrm{IC}}\frac{v_{\mathrm{pec},m}\,d}{d_{\mathrm{GA}}} \quad \text{for } d \ll d_{\mathrm{GA}} \tag{30}$$

d_{GA} is the distance to the GA while θ is the angle on our sky between it and the target galaxy, which is at a heliocentric distance d. The parameter $\beta_{IC} \approx 0.76$ accounts for initial condition drag.

We assume that the GA caused the LG to gain a peculiar velocity of $v_{\mathrm{pec,LG}} = 630\ \mathrm{km\ s^{-1}}$. The direction towards the GA is taken to be $l = 325^\circ$, $b = -7^\circ$ in Galactic coordinates (Kraan-Korteweg 2000). This is similar to the direction of $\boldsymbol{v}_{\mathrm{pec,LG}}$. The discrepancy is likely caused by less massive objects closer to the LG. For example, Centaurus A, M81, IC 342 and the VC are all in the northern Galactic hemisphere, partly explaining why $\boldsymbol{v}_{\mathrm{pec,LG}}$ points further north than the direction towards the GA.

To estimate β_{IC}, we construct a basic simulation involving a particle of mass M in an otherwise homogeneous Universe. A test particle some distance r away is required to satisfy $r = d$ with a peculiar velocity close to 630 km s^{-1} at the present time. This is achieved by varying its position when the cosmic scale-factor $a = 0.1$, at which time the particle satisfied equation (1). In section 2.1 of Banik & Zhao (2016), we used general relativity to show that the test particle satisfies

$$\ddot{r} = \frac{\ddot{a}}{a}r - \frac{GM}{r^2} \tag{31}$$

We solve this problem using a Newton–Raphson procedure targeting a present distance of $d = 84$ Mpc. We then repeat the calculation with a slightly different d. Without the GA, the relative RV of the particles would be $H_0 r$. Thus, we determine the effect of the GA using

$$\Delta\mathrm{GRV}_{\mathrm{GA}} = \Delta v_{\mathrm{pec}} \quad \text{where} \tag{32}$$

$$v_{\mathrm{pec}} \equiv \dot{r} - H_0 r \tag{33}$$

In this way, we find that $\beta_{\mathrm{IC}} \approx 0.76$ for both cosmological models given in Table 2. Assuming $\Delta\mathrm{GRV}_{\mathrm{GA}}$ has the same angular dependence as the tidal field that causes it, we suppose that this calculation is sufficient to determine the result for all angles θ. This should be valid as long as the tides raised by the GA can be approximated as linear in position, which is reasonable within the LG.

For θ close to 0 or 180°, the GA tends to increase GRVs. However, for θ close to 90°, the GA reduces GRVs. This is because both the MW and the target galaxy accelerate towards the perturber at similar rates. As their comoving distance from the GA decreases, so also does their comoving distance from each other. Thus, the GA can only increase GRVs along one direction. It must reduce them along the other two, albeit by half as much. This is due to the divergence-free nature of the gravitational field far from its source. As a result, an external tidal field on the LG does not readily resolve the high-velocity galaxy problem within it as these galaxies lie in several quite different sky directions (Table 4).

None the less, one can hope that the GA helps with some of the most problematic cases. To see if this is likely, we compare the effect of the GA predicted by equation (30) and the ΔGRV of each galaxy. This is shown in Fig. 11. Here, we have propagated distance uncertainties in the usual way. The error budget on each ΔGRV accounts for uncertainties in the corresponding HRV and distance measurements (equation 10). A different GA distance to the assumed 84 Mpc causes a rescaling of its effects.

Equation (30) suggests that the GA should actually have *reduced* the GRV of most or all of the galaxies with the highest ΔGRV

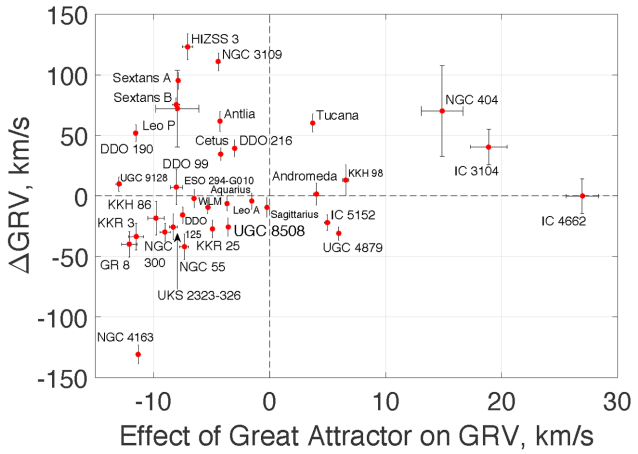


Figure 11. ΔGRVs are shown against our estimate for how much the Great Attractor might have increased the GRV of each galaxy (equation 30). Errors are correlated because a larger distance increases the effect of the GA while reducing ΔGRV.

(listed in Table 4). This remains the case if Antlia is treated as a low-mass satellite of NGC 3109 and excluded from our sample (van den Bergh 1999). Thus, far from helping to resolve the high-velocity galaxy problem, consideration of the GA appears to make it worse.

Only one galaxy with $\Delta\mathrm{GRV} < 0$ has a GRV so low as to be problematic given the expected ∼25 km s^{-1} accuracy of our model (Banik & Zhao 2016, Table 3). This is NGC 4163, which we have discussed previously. Our results suggest that its GRV would be reduced by the GA, thus helping to explain its low observed GRV. However, the effect is almost an order of magnitude too small.

It is straightforward to consider all the LG target galaxies in our sample, not just those substantially discrepant with our best-fitting model. After adjusting GRV predictions for the GA using equation (30) with different assumptions regarding the parameter α (equation 23), we found that the rms ΔGRV increased very slightly.

$$\text{Increase in rms } \Delta\mathrm{GRV} \approx (0.84 - 0.36\alpha)\ \mathrm{km\,s^{-1}} \tag{34}$$

The very small effect of the GA on the rms ΔGRV could well be a sign that our algorithm has utilized its many degrees of freedom to adjust the tidal field acting on the LG so as to best match observations.

Taken at face value, our analysis suggests that even galaxies at the edge of the LG ∼3 Mpc away would have only $\Delta\mathrm{GRV}_{\mathrm{GA}} \sim 35$ km s^{-1} at most. Typical effects would be $\sim\frac{1}{4}$ as much due to smaller distances and projection effects. Considering that several galaxies have a much larger ΔGRV such that even its rms value exceeds 40 km s^{-1}, the effect of the GA is too little to substantially alter our conclusions.

However, cosmological simulations based on ΛCDM suggest that structure tends to form along filaments (e.g. Springel et al. 2005, fig. 1). Thus, there may well be additional structures along the line of sight towards the GA (or on the opposite side) at smaller distances. The effect of these structures on the LG would have a similar angular dependence as that of the GA but possibly with an increased magnitude due to the smaller distance. This suggests that we may have underestimated the effects shown in Fig. 11.

Besides the GA, another structure outside our analysis that may be important to the LG is the VC. Using the same tidal field strength as for the GA, we estimated how much the VC could affect GRVs of objects in the LG (Fig. 12). In this case, the magnitude of the effect

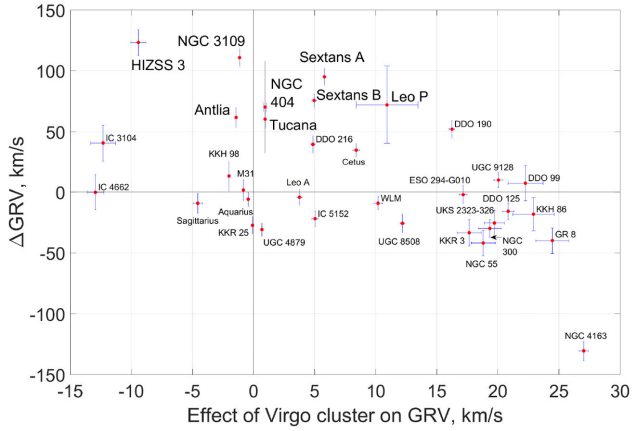


Figure 12. Similar to Fig. 11, for the Virgo Cluster. We use the same tidal field strength of 5.7 km s^{-1} Mpc^{-1} for target galaxies that make a right angle at the MW with the direction towards the Virgo Cluster ($l = 283.^{\circ}8$, $b = 74.^{\circ}4$, Nezri et al. 2012). Galaxies with a high ΔGRV have been emphasized.

is not known so well, making it more useful to focus on its sign as this depends only on the sky positions of the relevant objects. It is apparent that the VC may have some success at reducing the ΔGRVs of Sextans A, B and Leo P but would not much help with the other high ΔGRV galaxies. The overall trend is less promising as most galaxies are in the 'wrong' quadrants – either they have an anomalously high GRV but the VC would reduce their GRV, or vice versa. This is not the case with the GA, which would likely reduce GRVs of most LG galaxies (Fig. 11) and thus have little overall effect on how well our model matches observations (equation 34).

So far, we assumed that the GA and VC have always been in the same direction. Due to their large distances, they are almost certainly very close to the Hubble flow today. This was likely the case over the vast majority of the age of the Universe. It may not have been so at early times. However, because of initial condition drag, gravitational forces then have only a very small effect on present peculiar velocities (e.g. Banik & Zhao 2016, fig. 4). Thus, the crucial ingredient in predicting how the LG should be affected by, e.g. the GA is its present position.

5.2 Modified gravity

The high GRVs of some LG galaxies must have been caused by forces acting on them that our model does not account for. The nature of these forces might be better understood if we had an idea of the space–time location where they acted. Thus, it is necessary to estimate where these galaxies may have been. For the sake of clarity, we will focus on NGC 3109 because it has a high ΔGRV in our model (Table 4) and was also identified as having a rather high GRV in previous works (Teyssier et al. 2012; Pawlowski & McGaugh 2014).

In this regard, it is helpful to define comoving positions $\boldsymbol{x}$ that do not change for particles in a homogeneous Universe. Thus, they are related to physical positions $\boldsymbol{r}$ by

$$\boldsymbol{x} \equiv \frac{\boldsymbol{r}}{a\left(t\right)} \tag{35}$$

As we are dealing with unknown forces, we seek only a rough estimate of the comoving displacement d of NGC 3109. We assume that it is so distant as to be following the Hubble flow in our model, such that $d \approx 0$ in it. The actual value of d is estimated based on integrating the trajectory of NGC 3109 backwards in time, with the effect of gravity crudely included by using its ΔGRV instead of its observed peculiar velocity as a present boundary condition. For simplicity, we only consider motion along the line between the LG barycentre and the present position of NGC 3109. Thus, its present peculiar velocity (equation 33) for the purposes of this section is estimated as

$$v_{\rm pec,0} \approx \Delta\mathrm{GRV} \tag{36}$$

Due to the effect of Hubble drag (e.g. Banik & Zhao 2016, equation 24), the peculiar velocity of a free particle changes with time according to $v_{\rm pec} \propto a^{-1}$. Furthermore, even the same peculiar velocity corresponds to a more rapidly changing comoving position in the past (equation 35). Thus, integrating between some past time t_i and the present time t_0, we estimate that the comoving displacement

$$d = \int_{t_i}^{t_0} \frac{v_{\rm pec}}{a}\,\mathrm{d}t \tag{37}$$

$$= \int_{t_i}^{t_0} \frac{v_{\rm pec,0}}{a^2}\,\mathrm{d}t \tag{38}$$

For a rough estimate, it is acceptable to approximate the expansion history of the Universe as linear in time (equation 27). During the time period of interest ($a \gtrsim 0.2$), the results obtained in this way for $\int_{t_i}^{t_0} a^{-2}\,\mathrm{d}t$ are accurate to within ∼10 per cent. The error is much smaller if $a_i \equiv a\left(t_i\right)$ is close to 0.2 or 1. This allows us to solve equation (38).

$$d = \frac{v_{\rm pec,0}}{H_0}\left({a_i}^{-1} - 1\right) \tag{39}$$

$$= \frac{v_{\rm pec,0} z}{H_0} \quad \text{where} \tag{40}$$

$$z \equiv {a_i}^{-1} - 1 \tag{41}$$

To simplify our discussion, we assume that the high ΔGRV of NGC 3109 was caused by forces acting over a small fraction of the Hubble time, e.g. due to an encounter with a massive object. There is a trade-off between how long ago these forces acted and their total impulse. M31 is the fastest rotating LG galaxy, with a circular velocity of ∼225 km s^{-1} (Carignan et al. 2006). An impulse of twice this is possible only for an object not eventually accreted if it is on a circular orbit. Thus, we suppose that the impulse could not feasibly have exceeded triple the ∼110 km s^{-1} ΔGRV of NGC 3109, implying $a_i > \frac{1}{3}$. In the limiting case, an unexplained impulse of ∼330 km s^{-1} would have to be acquired when the redshift $z = 2$, presumably in a gravitational interaction.

Given that 110 km s^{-1} is very close to the Hubble flow rate at the distance of NGC 3109, we can set $v_{\rm pec,0} \approx H_0 d_0$, where the present comoving/physical distance from the LG barycentre to NGC 3109 is d_0. Thus, its comoving displacement d could not have much exceeded twice this, in which case the unexplained impulse occurred $\lesssim$1.7 comoving Mpc from the LG, possibly on the opposite side to where NGC 3109 currently lies. This makes it very difficult to understand how the present motion of NGC 3109 came about if one looks for an explanation outside the LG. Furthermore, the effect of gravity implies that galaxy groups outside it must have started at a slightly larger comoving separation with it than they presently have. For example, our model indicates that the comoving distance

between the MW–M31 mid-point[12] and Cen A has decreased from 4.98 Mpc to 4.04 Mpc since redshift 9.

A more plausible scenario might be that the MW and/or M31 is the massive object(s) responsible for the anomalous kinematics of NGC 3109. In this case, the missing ingredient is an impulse close to the LG barycentre, since which time the comoving displacement $d \approx d_0$. For this to occur, we need to set $a_i = \frac{1}{2}$ in equation (39), corresponding to ∼8 Gyr ago. The conclusion that NGC 3109 was likely close to the LG barycentre at this time has previously been reached using simpler methods (Pawlowski & McGaugh 2014).

However, it would be very unusual if the major LG galaxies were responsible for the high ΔGRV of NGC 3109. After all, our model directly includes the MW and M31 as well as their most massive satellites (Table 3). None the less, gravitational slingshot interactions with these objects could well lead to high GRVs, as occurs close to the LG barycentre (bottom panel of Fig. 1). Thus, increasing the efficiency of this process might help to explain the observations. The energy gained in such interactions is reliant on the gravitational potential of the massive body being time-dependent due to its motion. This suggests that the relative motion of the MW and M31 might have been much faster in the past than implied by our model. Given their known relative velocity at present, this implies a rather high mutual acceleration. Therefore, we need to consider whether we have correctly understood the gravitational effect of the MW and M31 on each other.

So far, our discussion has been restricted to models based on ΛCDM. We have seen that it faces difficulties in explaining the dynamics of LG galaxies, both using a 3D model and a thorough grid investigation of the parameters in an axisymmetric model (Banik & Zhao 2016).

Beyond the LG, some remarkably tight correlations exist between the dynamics of galaxies and the distribution of their luminous matter (e.g. Famaey & McGaugh 2012, and references therein). This has recently been confirmed and further tightened based on near-infrared observations by the *Spitzer Space Telescope* (McGaugh, Lelli & Schombert 2016). These correlations were unexpected in the context of ΛCDM. However, many of the trends were predicted a priori using Modified Newtonian Dynamics (MOND; Milgrom 1983). Thus, we consider whether this theory may shed light on the high-velocity galaxy problem. Our reasoning will be similar to that in section 4.6 of Banik & Zhao (2016).

MOND imposes an acceleration-dependent modification to the usual Poisson Equation of Newtonian gravity (Bekenstein & Milgrom 1984; Milgrom 2010). In spherical symmetry, the result is that the gravitational field g at distance r from an isolated point mass M transitions from the usual inverse square law at short range to

$$g = \frac{\sqrt{GMa_0}}{r} \quad \text{for } r \gg \frac{\sqrt{GM}}{a_0} \tag{42}$$

Here, a_0 is a fundamental acceleration scale of nature. Empirically, $a_0 \approx 1.2 \times 10^{-10}$ m s^{-2} to match galaxy rotation curves (McGaugh 2011). At this value, there is a remarkable coincidence with the acceleration at which the energy density in a classical gravitational field becomes comparable to the dark energy density u_Λ.[13]

Thus,

$$\frac{g^2}{8\pi G} < u_\Lambda c^2 \quad \Leftrightarrow \quad g \lesssim 2\pi a_0 \tag{43}$$

This suggests that MOND may be caused by quantum gravity effects (e.g. Milgrom 1999; Pazy 2013; Verlinde 2016). Regardless of the underlying microphysical explanation, at sufficiently low acceleration, MOND gravity from a point mass follows equation (42) as long as gravity from other objects is negligible.

The external gravitational field on the LG can be estimated based on its peculiar velocity of ∼630 km s^{-1} relative to the surface of last scattering (Planck Collaboration XXVII 2014). As might be expected, this shows the LG to be fairly isolated (Famaey, Bruneton & Zhao 2007). Thus, the force between the MW and M31 declines much slower with their separation than in ΛCDM, especially in the range between their virial radii (∼150 kpc in our models) and their actual separation (783 ± 25 kpc; McConnachie 2012).[14]

If correct, the much stronger force between these galaxies has dramatic consequences for the whole LG. This is because the MW–M31 orbit is almost radial (van der Marel et al. 2012b). As a result, MOND implies a past close flyby encounter between them 9 ± 2 Gyr ago (Zhao et al. 2013). The tidal tails expelled from the discs of these galaxies during their interaction may be responsible for the thin corotating system of satellites around the MW (e.g. Pawlowski & Kroupa 2013) and the similar system around M31 (Ibata et al. 2013). This interaction may also have formed the thick disc of the MW (Gilmore & Reid 1983), a structure that seems to have formed fairly rapidly from the thin disc 9 ± 1 Gyr ago (Quillen & Garnett 2001). More recent investigations also suggest a fairly rapid formation time-scale (Hayden et al. 2015). The disc heating that likely formed the thick disc appears to have been stronger in the outer parts of the MW, characteristic of a tidal effect (Banik 2014). This may be why it has a larger scale length than the thin disc of the MW (Jurić et al. 2008; Jayaraman et al. 2013).

At the point of closest approach, the relative velocity of the MW and M31 would have been ∼600 km s^{-1} (Zhao et al. 2013). Such fast motions could lead to very powerful gravitational slingshot encounters. The limiting factor might even have been their circular rotation velocities rather than the motions of their centres of mass. If we assume a maximum impulse of $\sim v_{f,M31}$ and an encounter when $a_i = \frac{1}{2}$, then it is easy to see why there are no galaxies with ΔGRV $\gtrsim$ 120 km s^{-1}.

For this explanation to work, the galaxies with a high GRV in our ΛCDM-based model need to have been flung out from close to the LG barycentre at around the time the MW and M31 had their interaction. This implies that galaxies with a higher ΔGRV should generally lie further away from the LG. Thus, it is interesting that the conclusions we reached above using NGC 3109 also hold with Tucana because its lower ΔGRV is compensated by a smaller distance from the LG. This can be seen visually if one draws a line through these galaxies on Fig. 13 and realizes that it passes close to the origin.

The ΔGRV $\propto d$ relation between velocities and distances d from the LG barycentre also seems to apply to HIZSS 3, Sextans A and Sextans B. In theory, it applies to Cetus and DDO 216, though their low values of ΔGRV may mean that this is just a coincidence. Antlia falls slightly below this relation, but its GRV might have been

[12] a reasonable estimate for their barycentre as our model prefers them to have nearly equal masses (Table 2).

[13] Dark energy is required to explain why $\ddot{a} > 0$ despite the attractive effect of gravity (Riess et al. 1998).

[14] At still greater distances, the force would transition to the usual inverse square law, but with a much higher normalization than in Newtonian gravity (Milgrom 1986).

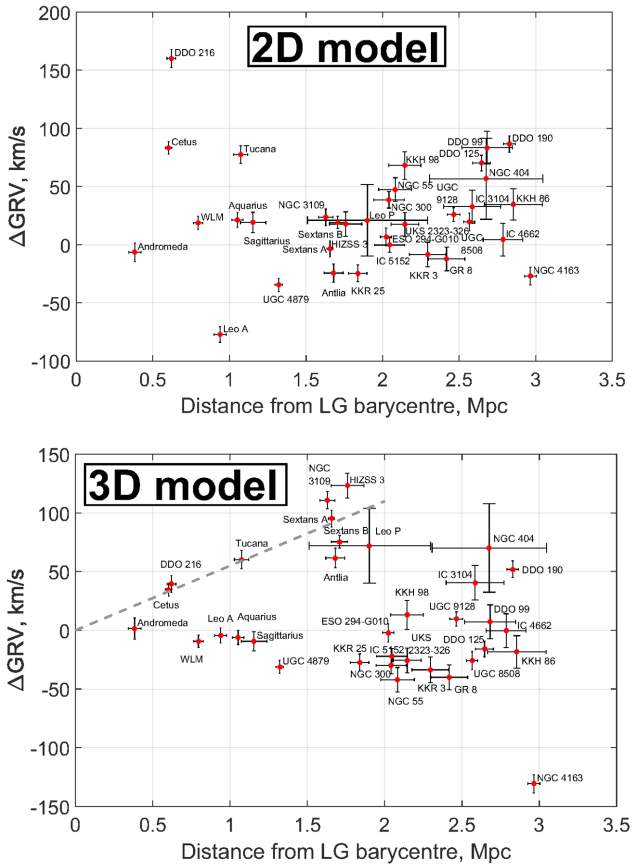


Figure 13. ΔGRV is shown for our target galaxies against their distance from the LG barycentre. Parameters of the models used are given in Table 2, with the best-fitting ones used for the relevant number of dimensions in the model. Errors shown tend to be anti-correlated because a larger distance to a target increases its predicted GRV, reducing its ΔGRV.

reduced due to the effects of NGC 3109. Tidal features in it suggest that the two may have already undergone a pericentre (Barnes & de Blok 2001). The distance to Leo P is still sufficiently uncertain that it is consistent with the $\Delta GRV \underset{\sim}{\propto} d$ relation.

NGC 404 appears to have just as large a ΔGRV as some of the galaxies just considered, despite being further from the LG. However, we showed in Section 5.1 that its GRV has likely been increased by ∼15 km s^{-1} due to the GA (Fig. 11) while the VC should not have affected its GRV much (Fig. 12). Thus, although the assumed 3.06 Mpc distance to NGC 404 seems to be correct (Dalcanton et al. 2009), its GRV is not that unusual in a ΛCDM context.

It is interesting that out of the 18 targets we have at distances of 2–3 Mpc from the LG barycentre, all of them are broadly consistent with expectations based on ΛCDM because none of them have a 1σ lower bound on their ΔGRV exceeding 60 km s^{-1}.[15] However, at least four galaxies like this exist at distances of 1–2 Mpc, despite having only 11 targets in this range. The hypothesis that such high-velocity galaxies are equally likely to exist in both distance bins can be ruled out at the 3σ confidence level. This suggests that the mechanism missing from our models can cause only high ΔGRVs out to ∼2 Mpc, quite unlike the effect of tides (which should be even stronger at greater distances). Consequently, we favour an explanation inside the LG and suggest that the crucial ingredient missing from our models is a past MW–M31 flyby.

[15] NGC 4163 has ΔGRV < -60 km s^{-1}, but we focus here on the frequency of galaxies with ΔGRV $\gg 0$ as the main problem with our models is the existence of several such objects, whereas there is only one object like NGC 4163 and it is anomalous on model-independent grounds (Makarov et al. 2013, table 2).

6 CONCLUSIONS

We construct axisymmetric and three-dimensional dynamical models of the LG in the standard ΛCDM cosmological model. Neither is able to provide a good match to the observed positions and velocities of galaxies within ∼1–3 Mpc of the LG barycentre (Fig. 7). This is despite our 3D model accounting for quite a large number of massive objects both within and outside the LG (Table 3) with fairly weak prior constraints on their masses (equation 18).

Both analyses reveal several galaxies with RVs much higher than model predictions. Galaxies with anomalously low RVs are rare (Figs 3 and 9). Thus, the high-velocity galaxy problem within the LG persists when using a 3D model for it. However, the particular galaxies that have anomalously high RVs are different in the 2D and 3D analyses, with Tucana being the only clear example of a galaxy with a high RV compared to the predictions of both models (Fig. 13).

We consider several possibilities for why there are so many LG galaxies with such high GRVs. Perhaps the most plausible in a ΛCDM context is tides raised by objects outside the LG (Section 5.1). The Great Attractor seems unable to reconcile the kinematics of these galaxies with our model (Fig. 11). The VC may help somewhat, though the overall trend is for it to raise the expected GRV of galaxies that already have an anomalously low GRV and vice versa (Fig. 12). Thus, we do not consider it likely that tides would help greatly to explain the unusually high GRVs of several LG galaxies. This is especially true when considering that our model has quite a lot of flexibility to adjust the tides raised on the LG by varying the masses of objects outside it (equation 18).

In the framework of ΛCDM, our axisymmetric and 3D results suggest that the past motions of the MW and M31 are too slow to explain the observed kinematics of LG galaxies. A similar challenge with high-velocity objects also exists in some systems far outside the LG. For example, the high relative velocity of the components of the Bullet Cluster (Tucker, Tananbaum & Remillard 1995) is difficult to reconcile with the gravity of their dark matter haloes acting over the age of the Universe (e.g. Thompson & Nagamine 2012; Kraljic & Sarkar 2015).

An explanation for the Bullet Cluster and other similar objects like El Gordo (Molnar & Broadhurst 2015; Ng et al. 2015) might well require a modification to our understanding of gravity on large scales. Indeed, cosmological *N*-body simulations in Modified Newtonian Dynamics (MOND, Milgrom 1983) could give rise to much higher pairwise velocities (Llinares, Zhao & Knebe 2009; Angus & Diaferio 2011).

Closer to home, MOND requires that the MW and M31 have undergone a past close flyby (Zhao et al. 2013). This is a consequence of their much stronger mutual gravitational attraction and the almost radial nature of their orbit (van der Marel et al. 2012b). Their higher relative velocity would likely help to explain observations of other LG galaxies through gravitational slingshot interactions with LG dwarfs. A past encounter between them could also have led to the formation of tidal dwarf galaxies, some of which might have ended up bound to neither and moving away from the LG at high speed.

Another interesting idea is that of dark matter as a superfluid (Berezhiani & Khoury 2016; Khoury 2016). In this model, phonons in this superfluid mediate forces between the baryons in a galaxy. This leads to MOND-like behaviour, helping to explain the observed tight correlation between the distribution of baryons in galaxies and their rotation curves (McGaugh et al. 2016).

However, galaxies still need to be surrounded by large (∼200 kpc) haloes of dark matter in the normal phase to account for weak lensing because the phonon-mediated force does not affect photon trajectories. Thus, the substantial galaxy–galaxy weak lensing signal (e.g. Brimioulle et al. 2013; Milgrom 2013) needs to be explained in much the same way as in ΛCDM. This means that interacting galaxies must experience strong dynamical friction between their dark matter haloes. As a result, the MW and M31 could never have approached closely in the context of this model because they would subsequently merge.

We suggest that such an interaction none the less occurred and would help to resolve the high-velocity galaxy problem in the LG. More work will be required to test this scenario.

ACKNOWLEDGEMENTS

The authors wish to thank P. J. E. Peebles for providing the algorithm used in this contribution. They also thank the referee and Stacy McGaugh for useful comments. IB is supported by a Science and Technology Facilities Council studentship. Results from the FORTRAN-based simulations were analysed and plotted using MATLAB®.

REFERENCES

Alcock C. et al., 2000, ApJ, 542, 281
Angus G. W., Diaferio A., 2011, MNRAS, 417, 941
Aragon-Calvo M. A., Silk J., Szalay A. S., 2011, MNRAS, 415, L16
Banik I., 2014, preprint (arXiv:1406.4538v2)
Banik I., Zhao H., 2016, MNRAS, 459, 2237
Barnes D. G., de Blok W. J. G., 2001, AJ, 122, 825
Begum A., Chengalur J. N., Karachentsev I. D., Sharina M. E., 2005, MNRAS, 359, L53
Bekenstein J., Milgrom M., 1984, ApJ, 286, 7
Bellazzini M., Oosterloo T., Fraternali F., Beccari G., 2013, A&A, 559, L11
Berezhiani L., Khoury J., 2016, Phys. Lett. B, 753, 639
Brimioulle F., Seitz S., Lerchster M., Bender R., Snigula J., 2013, MNRAS, 432, 1046
Brunthaler A., Reid M. J., Falcke H., Henkel C., Menten K. M., 2007, A&A, 462, 101
Carignan C., Chemin L., Huchtmeier W. K., Lockman F. J., 2006, ApJ, 641, L109
Carr B., 1994, ARA&A, 32, 531
Dalcanton J. J. et al., 2009, ApJS, 183, 67
Famaey B., McGaugh S. S., 2012, Living Rev. Relativ., 15, 10
Famaey B., Bruneton J.-P., Zhao H., 2007, MNRAS, 377, L79
Fermi-LAT Collaboration, 2015, Phys. Rev. Lett., 115, 231301
Francis C., Anderson E., 2014, Celest. Mech. Dyn. Astron., 118, 399
Gilmore G., Reid N., 1983, MNRAS, 202, 1025
Harris G. L. H., Rejkuba M., Harris W. E., 2010, PASA, 27, 457
Hayden M. R. et al., 2015, ApJ, 808, 132
Ibata R. A. et al., 2013, Nature, 493, 62
Jayaraman A., Gilmore G., Wyse R. F. G., Norris J. E., Belokurov V., 2013, MNRAS, 431, 930
Jurić M. et al., 2008, ApJ, 673, 864
Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2012, ApJ, 761, 98
Kahn F. D., Woltjer L., 1959, ApJ, 130, 705
Kallivayalil N., van der Marel R. P., Besla G., Anderson J., Alcock C., 2013, ApJ, 764, 161
Khoury J., 2016, Phys. Rev. D, 93, 103533
Kogut A. et al., 1993, ApJ, 419, 1
Komatsu E. et al., 2011, ApJS, 192, 18
Kraan-Korteweg R. C., 2000, in Page D., Hirsch J. G., eds, Lecture Notes in Physics Vol. 556, From the Sun to the Great Attractor. Springer, Berlin. p. 301
Kraljic D., Sarkar S., 2015, J. Cosmol. Astropart. Phys., 4, 050
Llinares C., Zhao H. S., Knebe A., 2009, ApJ, 695, L145
LUX Collaboration, 2017, Phys. Rev. Lett., 118, 021303
McConnachie A. W., 2012, AJ, 144, 4
McGaugh S. S., 2011, Phys. Rev. Lett., 106, 121303
McGaugh S., Lelli F., Schombert J., 2016, Phys. Rev. Lett., 117, 201101
McMillan P. J., 2011, MNRAS, 414, 2446
McMillan P. J., 2017, MNRAS, 465, 76
McQuinn K. B. W., Skillman E. D., Cannon J. M., Dalcanton J. J., Dolphin A., Stark D., Weisz D., 2009, ApJ, 695, 561
McQuinn K. B. W. et al., 2013, AJ, 146, 145
McQuinn K. B. W. et al., 2015, ApJ, 812, 158
Ma J., Wu Z., Wang S., Fan Z., Zhou X., Wu J., Jiang Z., Chen J., 2010, PASP, 122, 1164
Makarov D. I., Makarova L. N., Uklein R. I., 2013, Astrophys. Bull., 68, 125
Mieske S., Hilker M., Infante L., 2005, A&A, 438, 103
Milgrom M., 1983, ApJ, 270, 365
Milgrom M., 1986, ApJ, 302, 617
Milgrom M., 1999, Phys. Lett. A, 253, 273
Milgrom M., 2010, MNRAS, 403, 886
Milgrom M., 2013, Phys. Rev. Lett., 111, 041105
Molnar S. M., Broadhurst T., 2015, ApJ, 800, 37
Nezri E., White R., Combet C., Hinton J. A., Maurin D., Pointecouteau E., 2012, MNRAS, 425, 477
Ng K. Y., Dawson W. A., Wittman D., Jee M. J., Hughes J. P., Menanteau F., Sifón C., 2015, MNRAS, 453, 1531
Ostriker J. P., Peebles P. J. E., 1973, ApJ, 186, 467
PandaX-II Collaboration, 2016, Phys. Rev. Lett., 117, 121303
Pawlowski M. S., Kroupa P., 2013, MNRAS, 435, 2116
Pawlowski M. S., McGaugh S. S., 2014, MNRAS, 440, 908
Pazy E., 2013, Phys. Rev. D, 87, 084063
Peebles P. J. E., Tully R. B., 2013, preprint (arXiv:1302.6982)
Peebles P. J. E., Tully R. B., Shaya E. J., 2011, preprint (arXiv:1105.5596)
Peñarrubia J., Fattahi A., 2016, preprint (arXiv:1609.07147)
Peñarrubia J., Ma Y.-Z., Walker M. G., McConnachie A., 2014, MNRAS, 443, 2204
Peñarrubia J., Gómez F. A., Besla G., Erkal D., Ma Y.-Z., 2016, MNRAS, 456, L54
Penny S. J., Pimbblet K. A., Conselice C. J., Brown M. J. I., Grützbauch R., Floyd D. J. E., 2012, ApJ, 758, L32
Pietrzyński G. et al., 2013, Nature, 495, 76
Pimbblet K. A., Couch W. J., 2012, MNRAS, 419, 1153
Planck Collaboration XXVII, 2014, A&A, 571, A27
Planck Collaboration XIII, 2016, A&A, 594, A13
Quillen A. C., Garnett D. R., 2001, in Funes J. G., Corsini E. M., eds, ASP Conf. Ser. Vol. 230, Galaxy Disks and Disk Galaxies. Astron. Soc. Pac., San Francisco, p. 87
Riess A. G. et al., 1998, AJ, 116, 1009
Sandage A., 1986, ApJ, 307, 1
Schmidt K. H., 1958, Astron. Nachr., 284, 76
Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Shappee B. J., Stanek K. Z., 2011, ApJ, 733, 124
Slipher V. M., 1913, Lowell Obs. Bull., 2, 56
Soszyński I., Gieren W., Pietrzyński G., Bresolin F., Kudritzki R.-P., Storm J., 2006, ApJ, 648, 375
Springel V. et al., 2005, Nature, 435, 629
Steigman G., Turner M. S., 1985, Nucl. Phys. B, 253, 375
Teyssier M., Johnston K. V., Kuhlen M., 2012, MNRAS, 426, 1808
Thompson R., Nagamine K., 2012, MNRAS, 419, 3560
Tisserand P. et al., 2007, A&A, 469, 387
Tucker W. H., Tananbaum H., Remillard R. A., 1995, ApJ, 444, 532

Tully R. B. et al., 2013, AJ, 146, 86
van den Bergh S., 1999, ApJ, 517, L97
van der Marel R. P., Fardal M., Besla G., Beaton R. L., Sohn S. T., Anderson J., Brown T., Guhathakurta P., 2012a, ApJ, 753, 8
van der Marel R. P., Besla G., Cox T. J., Sohn S. T., Anderson J., 2012b, ApJ, 753, 9
Verlinde E. P., 2016, preprint (arXiv:1611.02269)
Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2009, ApJ, 707, 354
Zhao H., Famaey B., Lüghausen F., Kroupa P., 2013, A&A, 557, L3

This paper has been typeset from a TEX/LATEX file prepared by the author.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **473,** 4033–4054 (2018)
Advance Access publication 2017 October 9

doi:10.1093/mnras/stx2596

A plane of high-velocity galaxies across the Local Group

Indranil Banik★ and Hongsheng Zhao

Scottish Universities Physics Alliance, University of St Andrews, North Haugh, St Andrews, Fife KY16 9SS, UK

Accepted 2017 October 4. Received 2017 October 4; in original form 2017 January 21

ABSTRACT

We recently showed that several Local Group (LG) galaxies have much higher radial velocities (RVs) than predicted by a 3D dynamical model of the standard cosmological paradigm. Here, we show that six of these seven galaxies define a thin plane with root mean square thickness of only 101 kpc despite a widest extent of nearly 3 Mpc, much larger than the conventional virial radius of the Milky Way (MW) or M31. This plane passes within ∼70 kpc of the MW–M31 barycentre and is oriented so the MW–M31 line is inclined by 16° to it. We develop a toy model to constrain the scenario whereby a past MW–M31 flyby in Modified Newtonian Dynamics (MOND) forms tidal dwarf galaxies that settle into the recently discovered planes of satellites around the MW and M31. The scenario is viable only for a particular MW–M31 orbital plane. This roughly coincides with the plane of LG dwarfs with anomalously high RVs. Using a restricted *N*-body simulation of the LG in MOND, we show how the once fast-moving MW and M31 gravitationally slingshot test particles outwards at high speeds. The most distant such particles preferentially lie within the MW–M31 orbital plane, probably because the particles ending up with the highest RVs are those flung out almost parallel to the motion of the perturber. This suggests a dynamical reason for our finding of a similar trend in the real LG, something not easily explained as a chance alignment of galaxies with an isotropic or mildly flattened distribution (probability = 0.0015).

Key words: gravitation – Galaxy: disc – Galaxy: kinematics and dynamics – galaxies: distances and redshifts – Local Group – dark matter.

1 INTRODUCTION

The standard cosmological paradigm (Λcold dark matter, ΛCDM) faces several challenges in the relatively well-observed Local Group (LG). In particular, the satellite systems of its two major galaxies – the Milky Way (MW) and Andromeda (M31) – both exhibit an unusual degree of anisotropy. For the MW, this has been suspected for several decades (Lynden-Bell 1976, 1982), though recent observations have greatly clarified the situation and its apparent tension with ΛCDM (Kroupa, Theis & Boily 2005). Proper motion measurements show that most of its satellites corotate within a well-defined plane (Pawlowski & Kroupa 2013). Moreover, recently discovered ultra-faint satellites, globular clusters and tidal streams independently prefer a similarly oriented plane (Pawlowski, Pflamm-Altenburg & Kroupa 2012). Although some flattening is expected in ΛCDM (e.g. Butsky et al. 2016), it remains difficult to explain the very small thickness of the MW satellite system and its coherent rotation (Pawlowski et al. 2015).

An analogous situation was suspected around M31 (Metz, Kroupa & Jerjen 2007, 2009). This has recently been confirmed by Ibata et al. (2013) using the Pan-Andromeda Archaeological Survey (McConnachie et al. 2009). Despite its greater distance, the detection of this highly anisotropic system is rather secure because it is almost edge-on as viewed from our perspective. A redshift gradient across it strongly suggests that it too is corotating (Ibata et al. 2013). Like the MW satellite system, it is difficult for ΛCDM to explain the observed properties of the M31 satellite system (Ibata et al. 2014).

Beyond the LG, satellite planes likely also exist around Centaurus A (Müller et al. 2016) and M81 (Chiboucas et al. 2013). These discoveries may be related to a vast plane of dwarf galaxies recently found near M101 (Müller et al. 2017). This does not consist solely of satellite galaxies because it extends over 3 Mpc. Such a structure appears difficult to find in cosmological simulations of the ΛCDM paradigm (González & Padilla 2010, fig. 8).

We recently uncovered another potential problem relating to the dynamics of non-satellite LG dwarf galaxies at distances of ∼1–3 Mpc (Banik & Zhao 2016). This was based on a timing argument analysis of the LG (Kahn & Woltjer 1959; Einasto & Lynden-Bell 1982) extended to include test particles representing LG dwarfs. Following on from previous spherically symmetric dynamical models (Sandage 1986; Peñarrubia et al. 2014), we constructed an axisymmetric model of the LG consistent with the almost radial MW–M31 orbit (van der Marel et al. 2012b) and the close

★E-mail: ib45@st-andrews.ac.uk

alignment of Centaurus A with this line (Ma et al. 1998). Treating LG dwarfs as test particles in the gravitational field of these three massive moving objects, we investigated a wide range of model parameters using a full grid search. None of the models produced a good fit, even when we made reasonable allowance for inaccuracies in our model as a representation of ΛCDM based on the scatter about the Hubble flow in detailed *N*-body simulations of it (Aragon-Calvo, Silk & Szalay 2011). This is because several LG dwarfs have Galactocentric radial velocities (GRVs) much higher than expected in our best-fitting model, though the opposite was rarely the case (Banik & Zhao 2016, fig. 9). We found that this should remain true even when certain factors beyond the model are included, in particular the Large Magellanic Cloud and the Great Attractor (GA).

We borrowed an algorithm described in Peebles, Tully & Shaya (2011) to test whether this remains the case when using a 3D model of the LG. The typical mismatch between observed and predicted GRVs in the best-fitting model is actually slightly higher than in the 2D case, with a clear tendency persisting for faster outward motion than expected (Banik & Zhao 2017, figs 7 and 9). These results are similar to those obtained by Peebles (2017) using a similar algorithm. Despite a very different method to Banik & Zhao (2016), the conclusions remain broadly similar.

Beyond the LG, another puzzling observation in a ΛCDM context is the remarkably tight correlation between the internal accelerations within galaxies (typically inferred from their rotation curves) and the prediction of Newtonian gravity applied to the distribution of their luminous matter (e.g. Famaey & McGaugh 2012, and references therein). This 'radial acceleration relation' (RAR) is a generalization of the baryonic Tully–Fisher relation (e.g. McGaugh & Schombert 2015) which only considers the flat outer part of galaxy rotation curves and their total baryonic masses (equation 3), whereas the RAR considers all radii with accurate data.

The RAR has recently been confirmed and further tightened based on near-infrared photometry taken by the *Spitzer Space Telescope* (Lelli, McGaugh & Schombert 2016), considering only the most reliable rotation curves (as described in its Section 3.2.2 and also in Swaters et al. 2009) and taking advantage of reduced variability in stellar mass-to-light ratios at these wavelengths (Bell & de Jong 2001; Norris et al. 2016). These improvements reveal that the RAR holds with very little scatter over ∼5 orders of magnitude in luminosity and a similar range of surface brightness (McGaugh, Lelli & Schombert 2016).

In addition to disc galaxies, the RAR also seems to hold for ellipticals, whose internal forces can sometimes be measured accurately due to the presence of a thin rotation-supported gas disc (den Heijer et al. 2015). As well as these massive ellipticals, the RAR also works well in galaxies as faint as the satellites of M31 (McGaugh & Milgrom 2013). For a recent overview of how well the RAR works in several different types of galaxy across the Hubble sequence, we refer the reader to Lelli et al. (2017).

The RAR is either a fundamental consequence of natural law or an emergent property of galaxies relating their baryonic and dark matter distributions. The latter approach is taken by ΛCDM, a paradigm in which a relation of this sort is expected because lower mass dark matter haloes have shallower gravitational potential wells. This should make it easier for baryons to be ejected via energetic processes like supernova feedback. Still, the tightness of the observed RAR is difficult to explain in this way (Desmond 2017). Some attempts have been made to do so (e.g. Keller & Wadsley 2017), but so far these have investigated only a very small range of galaxy masses and types. In these limited circumstances, there does seem to be a tight correlation of the sort observed. However, a closer look reveals that several other aspects of the simulations are inconsistent with observations (Milgrom 2016). For example, the rotation curve amplitudes are significantly overestimated in the central regions (Keller, Wadsley & Couchman 2016, fig. 4).

Unlike ΛCDM, Modified Newtonian Dynamics (MOND, Milgrom 1983) is predicated on the assumption that the RAR is fundamental and not due to galaxies being surrounded by dark matter haloes. The dynamical effect of these haloes is instead provided by a revised law of gravity arising from an acceleration-dependent modification to the Poisson equation of Newtonian gravity (Bekenstein & Milgrom 1984; Milgrom 2010). In spherical symmetry, the gravitational field strength g at distance r from an isolated point mass M transitions from the usual inverse square law at short range to

$$g = \frac{\sqrt{GMa_0}}{r} \quad \text{for} \quad r \gg \sqrt{\frac{GM}{a_0}} \tag{1}$$

Here, a_0 is a fundamental acceleration scale of nature. Empirically, $a_0 \approx 1.2 \times 10^{-10}$ m s^{-2} to match galaxy rotation curves (McGaugh 2011). Remarkably, this is similar to the acceleration at which the energy density in a classical gravitational field becomes comparable to the dark energy density $u_\Lambda = \rho_\Lambda c^2$ implied by the accelerating expansion of the Universe (Riess et al. 1998). Thus,

$$\frac{g^2}{8\pi G} < u_\Lambda \;\Leftrightarrow\; g \lesssim 2\pi a_0 \tag{2}$$

This suggests that MOND may arise from quantum gravity effects (e.g. Milgrom 1999; Pazy 2013; Verlinde 2016; Smolin 2017). Regardless of its underlying microphysical explanation, MOND can explain the Tully–Fisher relation (Tully & Fisher 1977) as a specific example of the RAR by equating the gravitational field strength given by equation (1) with the centripetal acceleration $\frac{v^2}{r}$ required to maintain a circular orbit. In the low-acceleration outskirts of galaxies beyond the extent of most of their visible mass,[1] this predicts a flat rotation curve with amplitude

$$v_f = \sqrt[4]{GMa_0} \tag{3}$$

Although this is one of the more widely known consequences of MOND, the theory does much more than this and more even than the RAR, its prediction in isolated systems. For the recently discovered Crater 2 satellite of the MW (Torrealba et al. 2016), it predicted the velocity dispersion to be a tiny 2.1 km s^{-1} (McGaugh 2016b), partly due to a unique effect in MOND whereby its self-gravity is weakened by the external gravitational field of the nearby MW (e.g. Banik & Zhao 2015). This was recently confirmed by observations, which are in tension with a naive application of the RAR but not a more rigorous treatment of MOND (Caldwell et al. 2017). This external field effect hardly matters for calculating the rotation curve of the MW but is crucial to its escape velocity, measurements of which can be fit reasonably well in MOND (Banik & Zhao 2018).

A crucial ingredient for the RAR is the strength of the gravitational field in the outskirts of galaxies. These are impossible to measure directly and can only be estimated from rotation curves. Gravitational lensing provides an independent way to check these estimates in a statistical sense. One such attempt was the Canada–France–Hawaii Telescope Lensing Survey (Brimioulle et al. 2013). Stacked data from it show that MOND can predict the correct amplitude of weak gravitational lensing by spiral and elliptical galaxies, using equation (1) with the same value for a_0 as that required to

[1] Where a point mass approximation to the galaxy should be valid.

match disc galaxy rotation curves (Milgrom 2013). Thus, weak lensing and rotation curve measurements broadly agree on the strength of gravity in the outskirts of galaxies.[2]

As well as affecting forces within a galaxy, MOND also affects forces between them. In the LG, this implies a much stronger MW–M31 mutual attraction than ΛCDM. Combined with the almost radial nature of their relative motion (van der Marel et al. 2012b), this means that they must have undergone a close encounter $\sim 9 \pm 2$ Gyr ago (Zhao et al. 2013). This could have led to the formation of a thin tidal tail which later condensed into satellite galaxies of the MW and M31, a phenomenon which seems to occur in some observed galactic interactions (Mirabel, Dottori & Lutz 1992) and in MOND simulations of them (Tiret & Combes 2008).The formation mechanism of these tidal dwarf galaxies would lead to them lying close to a plane and corotating within that plane (Wetzstein, Naab & Burkert 2007), though a small fraction might well end up counter-rotating (Pawlowski, Kroupa & de Boer 2011). Some could even become unbound from both the MW and M31, instead flying away from the LG at high speed. This is possible once the effect of dark energy is considered as its repulsive effect rises with distance, unlike the gravitational field from a finite distribution of matter (equation 22).

A past MW–M31 interaction might also have formed the thick disc of the MW (Gilmore & Reid 1983), a structure which seems to have formed fairly rapidly from its thin disc 9 ± 1 Gyr ago (Quillen & Garnett 2001). More recent investigations suggest a fairly rapid formation time-scale (Hayden et al. 2015) and an associated burst of star formation (Snaith et al. 2014, fig. 2). The disc heating which likely formed the Galactic thick disc appears to have been stronger in the outer parts of the MW, characteristic of a tidal effect (Banik 2014). This may be why the thick disc of the MW has a larger scalelength than its thin disc (Jurić et al. 2008; Jayaraman et al. 2013).

The high MW–M31 relative velocity around the time of their encounter ($\sim 600\,\mathrm{km\,s^{-1}}$, Zhao et al. 2013) suggests that they could well have flung out several LG dwarfs at high speed in what would essentially have been three-body gravitational interactions (Banik & Zhao 2016). The main objective of the present contribution is to test certain aspects of this scenario. In Section 2, we extract some of its likely consequences based on a toy model of a past MW–M31 flyby encounter. Linking this to the observed geometry of the LG gives a constraint on the MW–M31 orbital plane. In Section 3, we use this in a more detailed MOND simulation of the LG incorporating several hundred thousand test particles affected by the gravity of the MW and M31, which undergo a close (14.17 kpc) flyby 6.59 Gyr after the big bang. As expected, some particles are flung out at high RVs after passing close to the space–time location of this event. The particles flung out to the greatest distances have orbital angular momenta aligning rather closely with that of the MW–M31 orbit (Fig. 8) and lie rather close to the MW–M31 orbital plane (Fig. 9). This is probably because such particles were ejected almost parallel to the motion of the perturbing body in order to gain the most energy from it.

In Section 4.1, we refine our previous ΛCDM model of the LG in 3D (Banik & Zhao 2017) to help us better select LG galaxies whose kinematics suggest that they were flung out in this way. We quantify the spatial anisotropy of these galaxies in Section 5. Here, we use our MOND-based simulation to identify three further properties that we expect of these high-velocity galaxies (HVGs). In Section 6, we quantify how likely it is for a random distribution of HVGs to match these properties as well as the observed HVG system. In Section 7, we discuss our analysis in light of previous works and consider some possibilities for explaining our results within MOND (Section 7.1) and ΛCDM (Section 7.2). Our conclusions are given in Section 8.

[2] Although MOND is a non-relativistic theory, all attempts to generalize it to the relativistic case imply that the non-relativistic gravitational field determines light deflection in the same way as in General Relativity (Milgrom 2013, section 2).

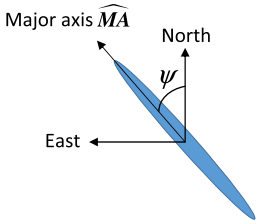


Figure 1. This is how an external disc galaxy-like M31 appears on our sky. The direction towards the centre of its image (into screen) is $\widehat{\boldsymbol{r}}_{\rm M31}$. Its position angle ψ is defined as the angle of its major axis $\widehat{\mathbf{MA}}$ eastwards of the local North (equation 5). Its inclination i to the sky plane can be determined from the ellipticity of its image.

Table 1. Observational parameters of M31 important for this work. Its sky position in Galactic coordinates (latitude last) is from Evans et al. (2010). We also give its inclination (Ma 2001) and position angle (Chemin, Carignan & Foster 2009, section 5.2). Combined with RV measurements, this implies its disc has a particular spin vector $\widehat{\boldsymbol{h}}_{\rm M31}$, which we find using equations (6) and (7). By convention, $\widehat{\boldsymbol{h}}_{\rm MW}$ points towards the South Galactic Pole.

Variable	Meaning	Value
$\widehat{\boldsymbol{r}}_{\rm M31}$	Direction to M31 now	(121.°57, −21.°57)
i	Inclination of M31 disc to sky plane	77.°5
ψ	Position angle of M31 disc on sky (Fig. 1)	$37.°7 \pm 0.°9$
$\widehat{\boldsymbol{h}}_{\rm M31}$	Internal angular momentum direction of M31 disc	(238.°65, −26.°89)

2 GEOMETRY OF A PAST MW–M31 FLYBY

2.1 Orientation of the M31 disc

We begin by describing how we find the angular momentum direction of the M31 disc $\widehat{\boldsymbol{h}}_{\rm M31}$, where we define the unit vector $\widehat{\boldsymbol{v}} \equiv \frac{\boldsymbol{v}}{|\boldsymbol{v}|}$ for any vector $\boldsymbol{v}$. Based on the ellipticity of its image, we know the inclination i of the M31 disc to the plane of our sky. The orientation of this image is described by a position angle ψ, whose meaning is illustrated in Fig. 1. Our adopted values for these parameters are given in Table 1, the caption of which contains the relevant references.

The major axis of the M31 image corresponds to the direction $\widehat{\mathbf{MA}} \propto \widehat{\boldsymbol{r}}_{\rm M31} \times \widehat{\boldsymbol{h}}_{\rm M31}$, which is orthogonal to both the direction $\widehat{\boldsymbol{r}}_{\rm M31}$ towards M31 and to $\widehat{\boldsymbol{h}}_{\rm M31}$ as it must lie within both the sky and M31 disc planes. This leaves two possible directions for $\widehat{\mathbf{MA}}$. The convention is to take the one most nearly pointing east, whose local direction $\widehat{\mathbf{E}}$ and that of the local north $\widehat{\mathbf{N}}$ are

$$\widehat{\mathbf{E}} = \frac{\widehat{\mathbf{NCP}} \times \widehat{\boldsymbol{r}}_{\rm M31}}{|\widehat{\mathbf{NCP}} \times \widehat{\boldsymbol{r}}_{M31}|} \text{ at M31 position} \tag{4}$$

$$\widehat{\mathbf{N}} = \widehat{\boldsymbol{r}}_{\rm M31} \times \widehat{\mathbf{E}} \tag{5}$$

$\widehat{\mathbf{E}}$ is orthogonal to $\widehat{\boldsymbol{r}}_{\rm M31}$ and to $\widehat{\mathbf{NCP}}$, the direction of the North Celestial Pole. Knowing $\widehat{\mathbf{E}}$ fixes the choice of $\widehat{\mathbf{MA}}$ because of the convention that $\widehat{\mathbf{MA}} \cdot \widehat{\mathbf{E}} \geq 0$. This allows us to determine the position angle of M31

$$\psi = \cos^{-1}(\widehat{\mathbf{N}} \cdot \widehat{\mathbf{MA}}), 0 \leq \psi < 180^\circ \tag{6}$$

The inclination of M31 is the angle of its disc normal to the line of sight towards its centre, $\widehat{\boldsymbol{r}}_{\rm M31}$. Thus,

$$i = \cos^{-1} |\widehat{\boldsymbol{r}}_{\rm M31} \cdot \widehat{\boldsymbol{h}}_{\rm M31}|, \quad 0 \leq i \leq 90^\circ \tag{7}$$

These constraints on i and ψ can be satisfied if we reverse the sense in which M31 rotates (i.e. $\widehat{\boldsymbol{h}}_{\rm M31} \rightarrow -\widehat{\boldsymbol{h}}_{\rm M31}$). Its actual sense of rotation must be determined observationally. In Galactic coordinates, the northern part of M31 is receding from us relative to its southern part, indicating that its angular momentum must point further east. Thus, the Galactic longitude of $\widehat{\boldsymbol{h}}_{\rm M31}$ must exceed that of M31 itself (by $<180^\circ$). Combined with the other constraints, this unambiguously determines $\widehat{\boldsymbol{h}}_{\rm M31}$.

We used the 2D Newton–Raphson algorithm to vary the Galactic latitude and longitude of $\widehat{\boldsymbol{h}}_{\rm M31}$ in an attempt to match the available observational constraints on i and ψ (Table 1). Starting from a guess in the correct hemisphere, our algorithm converged on the same solution as that in table 4 of Raychaudhury & Lynden-Bell (1989), providing an important cross-check. However, no explanation was given there for how $\widehat{\boldsymbol{h}}_{\rm M31}$ was derived or the assumed M31 position angle and disc inclination, both of which we use more recent measurements for.

2.2 The MW–M31 orbital plane

The satellites of the MW mostly lie within a thin plane and corotate within it (Pawlowski & Kroupa 2013). The same is true for M31 (Ibata et al. 2013).[3] We investigate the scenario where these satellite planes were formed by a past close encounter between the MW and M31. Such an encounter is inevitable in MOND (Zhao et al. 2013) but impossible in ΛCDM as dynamical friction between their dark matter haloes would cause a rapid subsequent merger (e.g. Privon et al. 2013). This difference between the theories may provide a basis for distinguishing between them (Kroupa 2015).

We use a simple toy model to constrain the MW–M31 orbital angular momentum direction $\widehat{\boldsymbol{h}}$ required by this scenario. Our model is based on two simplifying assumptions − the tidal torque exerted by M31 on the MW is assumed to act only at the time of their closest approach and only on the part of the MW closest to M31 at that time (and vice versa). We also assume that $\widehat{\boldsymbol{r}}_{\rm M31}$ has rotated by an angle $\phi \approx 125^\circ$ since that time (Belokurov et al. 2014, fig. 9). Our calculations suggest that the actual value is very likely within 6° of this.

By definition, the present direction towards M31 must be orthogonal to $\widehat{\boldsymbol{h}}$, constraining $\widehat{\boldsymbol{h}}$ to lie along a great circle. We measure position along this great circle using the angle θ measured southwards from the point on it in the northern Galactic hemisphere at a Galactic longitude of 180°.

In our model, the tidal torque exerted on galaxy i

$$\Delta \boldsymbol{h}_{\rm i} \propto (\widehat{\boldsymbol{h}}_{\rm i} \times \widehat{\boldsymbol{r}})(\widehat{\boldsymbol{h}}_{\rm i} \cdot \widehat{\boldsymbol{r}}) \text{ where } i = \text{ MW or M31} \tag{8}$$

Galaxy i has its disc angular momentum in the direction $\widehat{\boldsymbol{h}}_{\rm i}$, while $\widehat{\boldsymbol{r}}$ is the direction towards the other galaxy *at the time of their closest approach*. equation (8) shows that we do not expect there to be much tidal torque on a galaxy if $\widehat{\boldsymbol{r}}$ is either within its disc plane or along its disc normal. Although this is not totally accurate, it does suggest that such solutions would have difficulty in explaining the large amount of tidal torque required to create the satellite plane of at least one major LG galaxy given the significant observed disc–satellite plane misalignment for both the MW and M31 (Fig. 15). Thus, we only consider solutions satisfying

$$\cos 87^\circ < |\widehat{\boldsymbol{h}}_{\rm i} \cdot \widehat{\boldsymbol{r}}| < \cos 3^\circ \text{ for both } i = \text{ MW and M31} \tag{9}$$

The material which eventually forms the satellite plane around galaxy i has an angular momentum parallel to

$$\widehat{\boldsymbol{h}}_{\rm i} + (\tan \kappa_{\rm i})\, \widehat{\Delta \boldsymbol{h}}_{\rm i} \tag{10}$$

The parameter $\tan \kappa_{\rm i}$ governs the relative importance of the tidal torque on galaxy i and the angular momentum its spinning disc already possessed before the interaction. Because the unit vectors $\widehat{\Delta \boldsymbol{h}}_{\rm i}$ and $\widehat{\boldsymbol{h}}_{\rm i}$ are orthogonal, $\tan \kappa_{\rm i}$ determines the model-predicted angle $\kappa_{\rm i}$ between the orientations of the disc and dominant satellite plane of galaxy i (Fig. 2). Without a more detailed model, it is difficult to estimate this angle. We assume that its tangent is in the range (0.1–10) and allow it to be different for the MW and M31 due to their different masses, disc sizes and rotation speeds.

For every value of θ, we vary the rotation angle ϕ between 119° and 131°. Each time, we find the value of $\tan \kappa_{\rm MW}$ that minimizes the angle between the calculated and observed spin vector of the MW satellite plane. The same procedure is used for M31. We combine these angular differences in quadrature to obtain a χ^2 statistic, which we base on an allowance of 10° for both satellite planes. Our results are shown in Fig. 3.

Restricting to angles $\phi \leq 131^\circ$, we can obtain a solution with $\chi^2 = 5.0$ for $\theta = 75^\circ$ and $\phi = 131^\circ$. Although the resulting χ^2 is a little higher than 2, it is still quite acceptable. Our toy model is thus able to provide a plausible explanation for the origins of the MW and M31 satellite planes based on these galaxies having undergone a past close encounter. Naturally, we hope to refine our model in future, perhaps by using the RAyMOND algorithm (Candlish, Smith & Fellhauer 2015) or the publicly available Phantom of RAMSES algorithm (Lüghausen, Famaey & Kroupa 2015), both of which handle MOND explicitly by adapting the RAMSES algorithm (Teyssier 2002) . It is already possible to use the latter to simulate interacting disc galaxies in MOND (Thies, Kroupa & Famaey 2016).

We find that it is much more difficult to explain the orientation of the M31 satellite system than that of the MW. This may be related to the M31 satellite plane being inclined to its disc by $\sim 47^\circ$ (Ibata et al. 2013), whereas the MW satellite plane is almost polar with respect to its disc (Pawlowski & Kroupa 2013). This makes it more likely that the M31 satellite plane has precessed from its initial orientation (Fernando et al. 2017), especially as

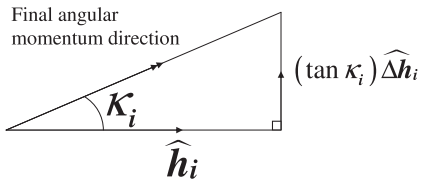


Figure 2. Illustration of how the MW–M31 interaction affects the angular momentum of material in the outer disc of galaxy i. In our model, the tidal torque on it is orthogonal to its original angular momentum $\widehat{\boldsymbol{h}}_{\rm i}$ arising from disc rotation. As a result, the disc–satellite plane mismatch angle $\kappa_{\rm i}$ measures the ratio of angular momentum gained to that originally present.

[3] Corotation can only be proved with proper motions, but it is strongly suggested by an RV gradient.

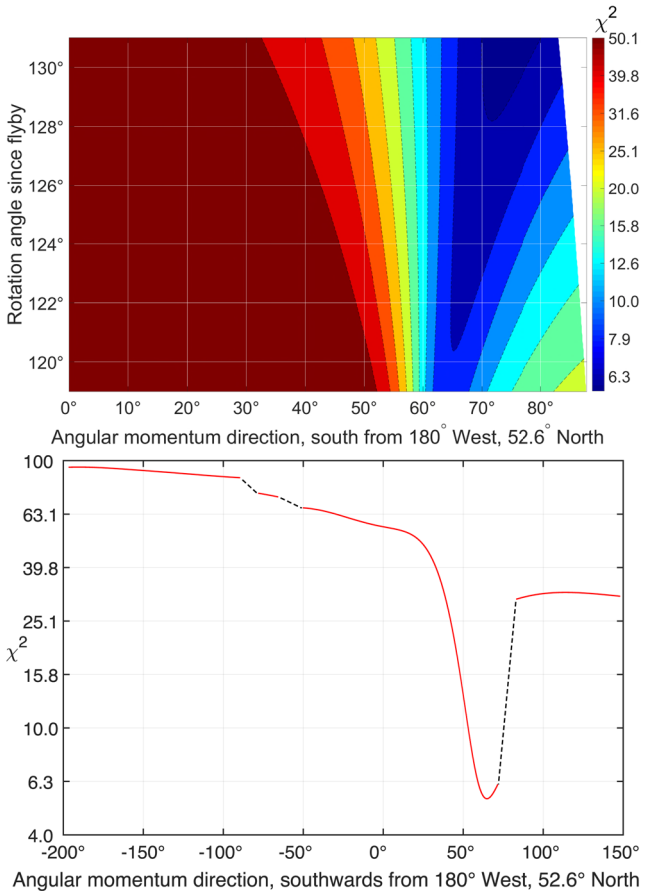


Figure 3. Top:goodness of fit χ^2 of our toy model to the MW and M31 satellite plane orientations as a function of model parameters, assuming an uncertainty of 10° for both systems. The best-fitting solution is given in Table 2. The gap arises as some models violate equation (9) and so could not plausibly lead to enough tidal torque on at least one major LG galaxy to explain the misalignment between its disc and dominant satellite plane (Fig. 15). The issue of most relevance is that the MW seems to have been almost within the M31 disc plane at the time of their closest approach. Bottom: χ^2 as a function of the adopted MW–M31 orbital plane with a fixed rotation angle of $\phi = 131°$ since their flyby. The dashed black lines correspond to directions of $\widehat{\boldsymbol{h}}$ for which it is not possible to satisfy the four constraints imposed by equation (9), leading to four distinct excluded ranges of θ (one of which has been folded around the figure for clarity).

the M31 disc has a scalelength ∼2.5× larger than that of the MW (Courteau et al. 2011; Bovy & Rix 2013). However, such precession effects would tend to thicken the M31 satellite plane as they would be weaker for satellites further from M31 (Fernando et al. 2017). The very low observed thickness (12.6 ± 0.6 kpc, Ibata et al. 2013) therefore argues against such an explanation.

Due to the different disc–satellite plane misalignments for the MW and M31, our model implies that the MW was more affected by tides from M31 than vice versa (Table 2). In MOND, we can obtain a good estimate of the mass of a galaxy from its rotation speed v_f in the flat outer region of its rotation curve (equation 3). This suggests that the slower rotating MW with $v_\mathrm{f} = 180\ \mathrm{km\,s^{-1}}$ (Kafle et al. 2012) has a lower mass than the faster rotating M31 with $v_\mathrm{f} = 225\ \mathrm{km\,s^{-1}}$ (Carignan et al. 2006). As the MW disc is also easier to tilt because it has less specific angular momentum than the M31 disc, it is reasonable to get $\tan\kappa_{MW}$ a few times larger than $\tan\kappa_{M31}$. Therefore, in MOND, the faster asymptotic rotation speed

Table 2. Values of the quantities most relevant to the geometry of a past MW–M31 flyby and its effects. The MW satellite plane orientation is from Pawlowski & Kroupa (2013, section 3) while that of M31 is from their section 4. Other parameters are determined using a toy model of a past MW–M31 interaction forming their planes of satellite galaxies (Section 2.2).

Quantity	Value
MW satellite plane spin vector	(176.°4, −15.°0)
M31 satellite plane spin vector	(206.°2, 7.°8)
MW–M31 orbital angular momentum direction θ	75°
Expected MW–M31 orbital pole $\widehat{\boldsymbol{h}}$	(217.°3, −15.°1)
Rotation angle of MW–M31 line since their flyby, ϕ	131°
$\tan\kappa_\mathrm{MW}$ (see Fig. 2)	3.7
$\tan\kappa_\mathrm{M31}$ (see Fig. 2)	1.0

of M31 is directly related to the larger observed angle between the MW disc and its satellite plane compared to the same quantity for M31 (Fig. 15).[4] This is exacerbated by our model implying that $\widehat{\boldsymbol{r}}_\mathrm{M31}$ at the time of the flyby lay only ∼6° out of the M31 disc plane, reducing how much torque could be exerted on it (equation 8). At that time, $\widehat{\boldsymbol{r}}_\mathrm{M31}$ was ∼66° from the MW disc plane. However, the larger scalelength of the M31 disc may have counteracted these factors somewhat. Combining these considerations suggests that

$$\frac{\tan\kappa_\mathrm{MW}}{\tan\kappa_\mathrm{M31}} \approx \left(\frac{v_\mathrm{f,MW}}{v_\mathrm{f,M31}}\right)^5 \frac{r_\mathrm{d,MW}}{r_\mathrm{d,M31}} \frac{\sin 66^\circ \cos 66^\circ}{\sin 6^\circ \cos 6^\circ} \tag{11}$$

$$\approx 4 \tag{12}$$

The ratio between the disc scalelengths r_d of the MW and M31 may have been different in the past and their orientations may have been slightly different too. Moreover, our model is only a very basic one. Despite this, equation (11) suggests that $\tan\kappa_\mathrm{MW} \approx 4\tan\kappa_\mathrm{M31}$, similar to the ratio in our best-fitting model (Table 2).

For both major LG galaxies, the rather large values of κ suggest that some material may have been pulled out of them and become unbound. This may explain why some LG non-satellite galaxies have unusual kinematics (Banik & Zhao 2017). It is possible that the material in some of them was once part of the MW or M31 disc or in their satellite system. Further study of this scenario must be left for future works.

3 THE LOCAL GROUP IN MOND

3.1 Governing equations

We now conduct a more detailed simulation of the LG in MOND, taking advantage of the MW–M31 orbital pole determined in Section 2.2. The MW–M31 trajectory is simulated by advancing them according to their mutual gravity supplemented by the cosmological acceleration term (e.g. Banik & Zhao 2016, equation 24).

$$\ddot{\boldsymbol{r}}_\mathrm{rel} = \boldsymbol{g}_\mathbf{M31} - \boldsymbol{g}_\mathbf{MW} + \frac{\ddot{a}}{a}\boldsymbol{r}_\mathrm{rel} \text{ where} \tag{13}$$

$$\boldsymbol{r}_\mathbf{rel} \equiv \boldsymbol{r}_\mathbf{M31} - \boldsymbol{r}_\mathbf{MW} \tag{14}$$

[4] These observational facts are otherwise unrelated.

Here, the cosmic scalefactor is $a(t)$ and $\boldsymbol{r}_\mathrm{i}$ is the position vector of galaxy i (MW or M31), at whose location the gravitational field (excluding self-gravity) is $\boldsymbol{g}_\mathrm{i}$. We use an overdot to indicate the time derivative of any quantity q e.g. $\dot{q} \equiv \frac{\partial q}{\partial t}$. All position vectors are with respect to the LG barycentre, which we take to be 0.3 of the way from M31 towards the MW. This is based on the asymptotic rotation curve of the MW flatlining at $\sim$180 km s^{-1} (Kafle et al. 2012), while the equivalent value for M31 is $\sim$225 km s^{-1} (Carignan et al. 2006). In the context of MOND, this suggests that the mass of M31 is $\left(\frac{225}{180}\right)^4 \approx 2.3\times$ that of the MW (equation 3).

Although MOND is known to work well at explaining the internal dynamics of galaxies outside the LG (e.g. Famaey & McGaugh 2012), we should check if this is the case for the MW and M31 before using it to determine the gravity they exert on each other and on the rest of the LG. For our neighbour M31, MOND can provide a fairly good match to its rotation curve using its observed baryonic distribution (Corbelli & Salucci 2007, fig. 4). For our work, it is important to note that this fit remains good out to rather large radii ($\sim$35 kpc or 7 disc scalelengths). A similar analysis for the MW is complicated slightly by our position within its disc. However, it has recently become clear that MOND can explain its rotation curve fairly well (McGaugh 2016a) and even provides a good match to its escape velocity curve (Banik & Zhao 2018). Thus, applying equation (3) to the MW and M31 rotation curves yields reasonable estimates for their baryonic masses, the vast majority of which resides in stars (91 per cent for M31 and 81 per cent for the MW, Yin et al. 2009, Section 2.2). This is almost certainly not representative of the Universe as a whole given that most of the mass in galaxy clusters is hot gas that has only recently been discovered at X-ray wavelengths (e.g. Vikhlinin et al. 2006). Indeed, the location of all the baryons in the Universe is far from certain, with significant amounts perhaps residing in an even more diffuse form (Nicastro, Mathur & Elvis 2008).

Having obtained MW and M31 masses (M_MW and M_M31) in this way, we treat them as point masses and find the gravitational field $\boldsymbol{g}$ they exert at position $\boldsymbol{r}$ using the quasi-linear formulation of MOND (Milgrom 2010). We assume the 'simple' interpolating function between the Newtonian and deep-MOND regimes (Famaey & Binney 2005) that works best for the MW rotation curve (Iocco, Pato & Bertone 2015).

$$\boldsymbol{g}_\mathrm{N} \equiv -\sum_{i=\mathrm{MW,M31}} \frac{GM_\mathrm{i}\left(\boldsymbol{r}-\boldsymbol{r}_\mathrm{i}\right)}{\left|\boldsymbol{r}-\boldsymbol{r}_\mathrm{i}\right|^3} \tag{15}$$

$$\nabla\cdot\boldsymbol{g} \equiv \nabla\cdot\left[\nu\left(\frac{|\boldsymbol{g}_\mathrm{N}|}{a_0}\right)\boldsymbol{g}_\mathrm{N}\right] \text{ where} \tag{16}$$

$$\nu\left(x\right) = \frac{1}{2}+\sqrt{\frac{1}{4}+\frac{1}{x}} \tag{17}$$

The appropriate boundary conditions are similar to Newtonian gravity, but for definiteness we give them here.

$$\nabla\times\boldsymbol{g} = 0 \tag{18}$$

$$\boldsymbol{g}\rightarrow 0 \text{ as } |\boldsymbol{r}|\rightarrow\infty \tag{19}$$

We use direct summation to obtain $\boldsymbol{g}$ from its divergence.

$$\boldsymbol{g}\left(\boldsymbol{r}\right) = \int \nabla\cdot\boldsymbol{g}\left(\boldsymbol{r}'\right)\frac{\left(\boldsymbol{r}-\boldsymbol{r}'\right)}{\left|\boldsymbol{r}-\boldsymbol{r}'\right|^3}\mathrm{d}^3\boldsymbol{r}' \tag{20}$$

$\nabla\cdot\boldsymbol{g}$is calculated out to almost 150$\times$ the MW–M31 separation, beyond which it should be very nearly spherically symmetric. Due to the shell theorem, it is unnecessary to consider $\nabla\cdot\boldsymbol{g}$ (or 'phantom dark matter') at even larger radii. As we only determine $\boldsymbol{g}$ out to 66.5$\times$ the MW–M31 separation, our results should be nearly free of edge effects. At larger distances, we assume that the MW and M31 can be treated as a single point mass located at their barycentre, yielding $\boldsymbol{g} = \nu\boldsymbol{g}_\mathrm{N}$.

3.2 MW–M31 trajectory

Using the gravitational field thus found, we integrate the MW–M31 trajectory backwards from present conditions. In general, the galaxies will not be on the Hubble flow at the start time of our simulations t_i, when the cosmic scalefactor $a_\mathrm{i} = 0.05$ and the Hubble parameter $H \equiv \frac{\dot{a}}{a}$ is H_i. However, deviations from the Hubble flow are observed to be very small at early times (Planck Collaboration XXVI, 2014). In order to satisfy this condition at t_i, we vary the total mass of the MW and M31 using a Newton–Raphson root-finding algorithm. This ensures that

$$\dot{\boldsymbol{r}}_\mathrm{rel} = H_\mathrm{i}\boldsymbol{r}_\mathbf{rel} \text{ when } t = t_\mathrm{i} \tag{21}$$

The MW and M31 are not on a purely radial orbit. Their mutual orbital angular momentum prevents them from converging on to the Hubble flow at very early times. This is unrealistic as any non-radial motion must have arisen due to tidal[5] torques well after the big bang. Thus, we take the MW–M31 orbit to be purely radial prior to their first turnaround at $t \approx 3$ Gyr. After this time, we assume their trajectory conserves angular momentum at its present value. This implies the MW–M31 angular momentum was gained near the time of their first turnaround, when their large separation would have strengthened tidal torques. At later times, the larger scalefactor would weaken tidal torques, suggesting that these have a much smaller effect around the time of the second MW–M31 turnaround than the first.

At present, there is no detailed theory for structure formation in MOND because it is unclear how to apply it to regions only slightly denser than the cosmic mean density. Assuming a particular model, structure formation was found to be more efficient than in ΛCDM (Llinares, Knebe & Zhao 2008). Observationally, there are several indications that this is actually the case (Peebles & Nusser 2010), including the rather high fraction of pure (bulgeless) disc galaxies (Kormendy et al. 2010).

Thus, structure formation in MOND – and perhaps in the Universe – is not so reliant on growth at late times through mergers, which would be less efficient in MOND due to the absence of dynamical friction between extended dark matter haloes. Instead, galaxies would form relatively rapidly after the big bang, emptying their surroundings due to the strong long-range gravitational attraction in the model. This would lead to many widely separated 'island universes' with fairly empty intervening voids, perhaps similar to the Local Volume (out to 8 Mpc) which does seem to contain voids emptier than might be expected in ΛCDM (Tikhonov & Klypin 2009). With mass draining on to a few well-separated massive galaxies, it would be natural for the MW and M31 to end up fairly isolated. Thus, there would not be much tidal torque on the MW–M31 system, leaving its orbit close to radial. This suggests that it would not be all that unusual for us to find ourselves in a galaxy which had a past close encounter with its nearest large neighbour if structure formation proceeded more efficiently than in ΛCDM. Of course, it remains to be seen whether it forms too efficiently in MOND.

[5] Affecting the MW and M31 differently.

Given the way we expect structure to form in MOND, we assume the MW and M31 masses do not grow by accretion at late times. However, an important effect included in our models is a 5 per cent reduction in their masses at the time of closest approach, when their simulated separation was just 14.2 kpc.[6] Considering that the MW disc has a scalelength of 2.15 kpc (Bovy & Rix 2013), while the corresponding quantity for M31 is 5.3 kpc (Courteau et al. 2011), it is very likely that some of the mass in these galaxies would be expelled to large distances and escape from them.

This mass could reside in the haloes of hot gas surrounding each galaxy. Such a halo has been detected around M31 based on absorption features in spectra of background quasars (Lehner, Howk & Wakker 2015). A similar halo is thought to be necessary around the MW to explain the truncation of the Large Magellanic Cloud's gas disc (Salem et al. 2015). These gas haloes seem to contain perhaps $3 \times 10^{10}\,\mathrm{M}_\odot$ each, with much larger amounts being very unlikely given constraints from the MW escape velocity curve (Banik & Zhao 2018). Considering that the MW rotation curve flatlines at $\sim$180 km s^{-1} (Kafle et al. 2012), while that of M31 flattens at $\sim$225 km s^{-1} (Carignan et al. 2006), MOND suggests their total baryonic mass is $2.3 \times 10^{11}\,\mathrm{M}_\odot$. This makes it quite feasible for them to have lost $10^{10}\,\mathrm{M}_\odot$ of hot gas around the time of their encounter, as our model implies. Some hot gas in an extended halo could also explain why the rotation curve-based estimate of the total MW and M31 mass falls a little below our timing argument estimate of $2.9 \times 10^{11}\,\mathrm{M}_\odot$ (for simplicity, we fix the MW:M31 mass ratio at 3:7 and scale up their masses slightly to make the timing argument work).

There are several other aspects of the problem which we include in our model using techniques we developed. We defer a more detailed explanation of our procedures to a forthcoming publication which will investigate the MW–M31 trajectory in MOND. For the present contribution, the major result is that equation (21) can be satisfied by backwards integration from present conditions using MW and M31 masses consistent with their rotation curves in MOND (themselves consistent with observed baryonic disc masses) and the more extended haloes of hot gas that have recently been detected around them (Lehner et al. 2015; Nicastro et al. 2016). The resulting MW–M31 trajectory is shown in Fig. 4. A past close encounter is inevitable in the context of MOND (Zhao et al. 2013) due to their slow relative tangential motion (van der Marel et al. 2012a) and the strong gravity in this model. We previously discussed how the thick disc of the MW and the LG satellite planes may well have formed due to this interaction (Sections 1 and 2.2, respectively). Here, we consider its effect on the rest of the LG.

3.3 Test particle trajectories

Once the MW–M31 trajectory is known, we can determine the gravitational field $\boldsymbol{g}$ everywhere within the LG at all times under the assumption that only these point masses are present in an otherwise homogeneous Universe. This allows us to advance the trajectories of test particles according to

$$\ddot{\boldsymbol{r}} = \frac{\ddot{a}}{a}\boldsymbol{r} + \boldsymbol{g} \tag{22}$$

$$\dot{\boldsymbol{r}} = H_\mathrm{i}\boldsymbol{r} \text{ when } t = t_\mathrm{i} \tag{23}$$

[6] This depends on the present proper motion of M31, for which only an upper limit is available (van der Marel et al. 2012a). Thus, the closest approach distance is unlikely to exceed 50 kpc but can be made arbitrarily small (Zhao et al. 2013).

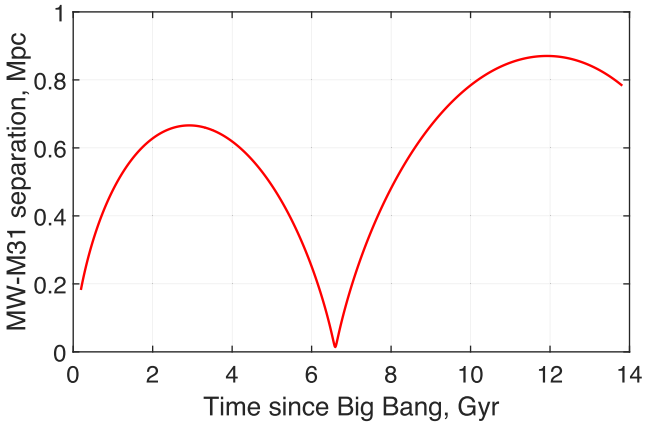


Figure 4. MW–M31 separation in our MOND simulation, showing a past close flyby 6.59 Gyr after the big bang at a closest approach distance of 14.17 kpc. At that time, their relative velocity was 716 km s^{-1}, of which 501 km s^{-1} was due to motion of the MW. The higher second apogalacticon is partly due to the effect of cosmology (equation 13) and our assumption that the MW and M31 lose 5 per cent of their mass around the time of their encounter (see the text).

Although the particles could not have started exactly on the Hubble flow, we consider this a reasonable assumption for reasons we now discuss. The MW and M31 could not have formed much earlier than when $a = 0.05$ because the MOND free-fall collapse time on to a point mass is given by

$$t_\mathrm{ff} = \frac{r_0}{v_\mathrm{f}}\sqrt{\frac{\pi}{2}} \tag{24}$$

v_f is given by equation (3) and taken to be 180 km s^{-1} for the MW (Kafle et al. 2012). Assuming the material currently in it must have turned around from a distance $\gtrsim$100 kpc (equation 27), this yields a free-fall time-scale of $t_\mathrm{ff} = 540$ Myr. This is much more than the 191 Myr age of the Universe when $a = 0.05$ in standard cosmology, suggesting that it is not appropriate to start our simulations much earlier.

Moreover, even if the MW existed as a point mass since the big bang, the resulting peculiar velocity some comoving distance d away would be

$$a v_\mathrm{pec} \equiv a\overbrace{(\dot{r} - Hr)}^{v_\mathrm{pec}} = \int_0^t a g \,\mathrm{d}t \text{ where} \tag{25}$$

$$g = \frac{\sqrt{GMa_0}}{ad} \tag{26}$$

The integrating factor $a(t)$ accounts for the effects of Hubble drag. We have assumed that the particle nearly follows the Hubble flow so that its distance to the mass can be taken as $a(t)d$. At a comoving distance of 2 Mpc, the peculiar velocity gained would only be 65 km s^{-1} by the time $a = 0.05$. At that time, the Hubble velocity of the particle was nearly 340 km s^{-1}, justifying our assumption that it was almost on the Hubble flow.

To understand the effect on the present-day velocity field, we must bear in mind that both the present position and velocity of a test particle would be affected if we get its velocity wrong at some earlier time. Thus, if we wish to know the velocity at a particular position today, we would need to consider a different test particle starting at a different position. This 'initial condition drag' scales down the effect of a velocity error at earlier times (when the scalefactor was a) on the present velocity *at fixed position* by a factor of $\sim a^{2.4}$ (Banik & Zhao 2016, fig. 4). However, even if we make the more

conservative assumption that it simply scales with a (like traditional Hubble drag), a 65 km s^{-1} velocity error when $a = 0.05$ would only affect the present LG velocity field by $\sim$3 km s^{-1}. This is probably why our axisymmetric dynamical model of the LG was hardly affected by using a different start time (Banik & Zhao 2016, section 4.6). Thus, our choice of initial conditions should be sufficient to get an approximate idea of how the LG might have been affected by a past MW–M31 flyby. Moreover, the fact that motions at high redshift have only a weak impact on our results implies that they should be robust to uncertainties surrounding the application of MOND in a cosmological context (at lower redshifts, the MW and M31 have already formed and so better approximate the isolated situations where it is clear how MOND works).

Because the MW and M31 must have accreted matter from some region prior to the start of our simulation, we exclude all test particles starting within a distance $r_{\rm exc,i}$ of galaxy i. We determine this by requiring that the excluded volume has as much baryonic matter as galaxy i, taking the density of baryons to be the cosmic mean value. This is obtained from the fraction $\Omega_{\rm b,\,0} = 0.049$ that baryons currently comprise of the cosmic critical density, which we found by taking $H_0 \equiv H\left(t_0\right) = 67.3\,{\rm km\,s^{-1}\ Mpc^{-1}}$ (Planck Collaboration XIII, 2016, table 4). The cosmic baryon density can be estimated using big bang nucleosynthesis arguments – only a narrow range of values is consistent with the primordial abundances of light elements such as deuterium (Cyburt et al. 2016).

$$\frac{4\pi}{3}{r_{\rm exc,i}}^3 \times \overbrace{\frac{3{H_0}^2}{8\pi G}\Omega_{\rm b,0}{a_{\rm i}}^{-3}}^{\text{Baryon density at } t_i} \equiv M_{\rm i}\left(\text{ for } r_{\rm exc,i}\right) \tag{27}$$

For consistency, it is necessary that the sizes of the excluded regions satisfy

$$r_{\rm exc,MW} + r_{\rm exc,M31} \leq |\boldsymbol{r}_{\rm rel}| \ \text{when} t = t_{\rm i} \tag{28}$$

This inequality applies because $r_{\rm exc}$ is 77.7 kpc for the MW and 102.5 kpc for M31, leading to a total of 180.2 kpc – interestingly, this is just smaller than $r_{\rm rel}\left(t_{\rm i}\right) = 182.1$ kpc, suggesting that the two galaxies accreted matter from regions which just touched. This remains the case if we use a slightly different start time as $r_{\rm exc,i} \propto a_{\rm i}$, similarly to $r_{\rm rel}\left(t_{\rm i}\right)$ – at such early times, both galaxies would follow the Hubble flow rather closely. However, this coincidence does not occur in ΛCDM, a model in which the excluded regions would very likely overlap (Banik & Zhao 2016, section 2.2.1).

3.4 Simulation results

In Fig. 5, we show the distances and RVs of test particles with respect to the LG barycentre, colour-coding them according to the orientation of their orbital plane. We quantify this based on the specific angular momentum $\boldsymbol{h}$, whose direction can readily be compared with the MW–M31 orbital pole $\widehat{\boldsymbol{h}}_{\rm MW-M31}$.

$$\boldsymbol{h} \equiv \boldsymbol{r} \times \dot{\boldsymbol{r}} \tag{29}$$

$$\cos\psi \equiv \widehat{\boldsymbol{h}} \cdot \widehat{\boldsymbol{h}}_{\rm MW-M31} \tag{30}$$

The initial positions of the particles span a grid in spherical polar coordinates. We do not show results for particles that pass within 15.4 kpc of the MW or 21.5 kpc of M31. In the real LG, such particles would likely have merged with the nearby galaxy.

For this work, the important feature is the upper branch of the Hubble diagram. Its upward slope arises because these particles must have passed close to the space–time location of the MW–M31 encounter and gained a substantial amount of kinetic energy in what was essentially a three-body interaction. Thus, for such particles to be further away from the LG now, they must have a larger outwards velocity. This will depend somewhat on when each particle approached whichever of the MW or M31 most strongly scattered that particle – and thus the speed of the perturber at that time.

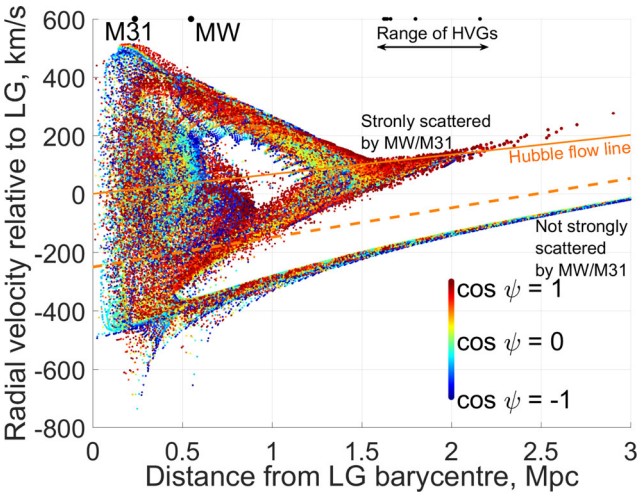


Figure 5. Hubble diagram of the test particles in our simulation coloured by their value of $\cos\psi$, which parametrizes how well their orbital angular momenta align with that of the MW–M31 orbit (equation 30). We show the Hubble flow line (solid orange) and a 1.5× steeper line (dashed orange) which we use to select analogues of HVGs in later figures. Particles below this line have generally never interacted closely with the MW or M31, unlike particles above the line. The black dots along the top edge of the figure indicate distances to the MW, M31 and the HVGs. Marker sizes have been enlarged in some regions for clarity, but do not correspond to volume factors (see the text). Thus, particles at distances $\lesssim$1 Mpc represent much less mass than it might appear.

However, the precise encounter distance b should not affect our results much because of the r^{-1} gravity law (equation 1). Roughly speaking, doubling b halves the strength of gravity but doubles the duration of the encounter, leaving the total impulse unchanged. This is probably why we found no clear correlation between how far particles were flung from the LG and how closely they approached the MW/M31. Thus, our results should not depend much on the minimum allowed encounter distance in our simulation or on the fact that the MW and M31 have finite extents and are not point masses. The important effect here is the time dependence of the gravitational field at the positions of the test particles which get flung out at high speed. Most of these never come that close to the MW or M31 because the commoner more distant encounters are just as effective at scattering test particles (Fig. 6).

ΛCDM also allows slingshot encounters with the MW and M31, but their fairly slow motion means that they can only fling galaxies out to $\sim$1 Mpc from the LG, at which point the upper branch of the Hubble diagram simply stops (Banik & Zhao 2016, fig. 3). Even in more detailed cosmological simulations of ΛCDM that include encounters with satellites of MW and M31 analogues, dwarf galaxies do not get flung out beyond this distance (Sales et al. 2007, figs 3 and 6). For MOND, the corresponding limit is $\sim$2.5 Mpc due to the MW–M31 flyby, which therefore makes a dramatic difference to the Hubble diagram at distances of $\sim$1–2 Mpc (Fig. 5). In this distance range, our simulation yields a bimodal distribution of RVs, with the HVGs corresponding to particles in the upper branch.

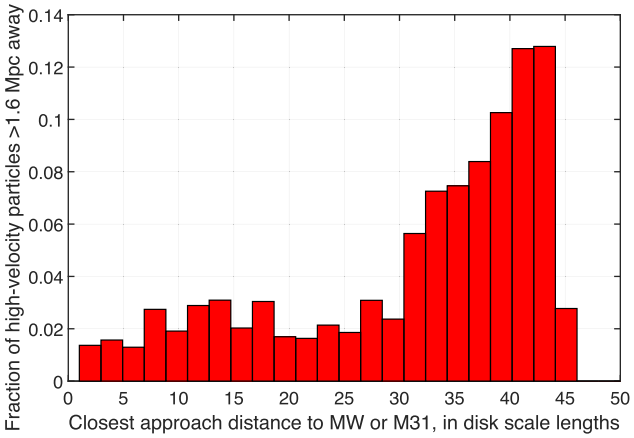


Figure 6. Histogram of the closest approach distances of high-velocity test particles (above dashed orange line in Fig. 5) to the MW or M31 in units of their disc scalelengths. We show whichever of these quantities is smaller for any given test particle. It is evident that very few of these particles approached the MW or M31 so closely for the details of their mass distribution to become important. Each particle has been weighted according to the volume it represents in our 3D grid of initial conditions. This corresponds to the mass it represents if we assume that the LG initially had a uniform density except in appropriately sized spherical 'holes' around the MW and M31 (equation 27).

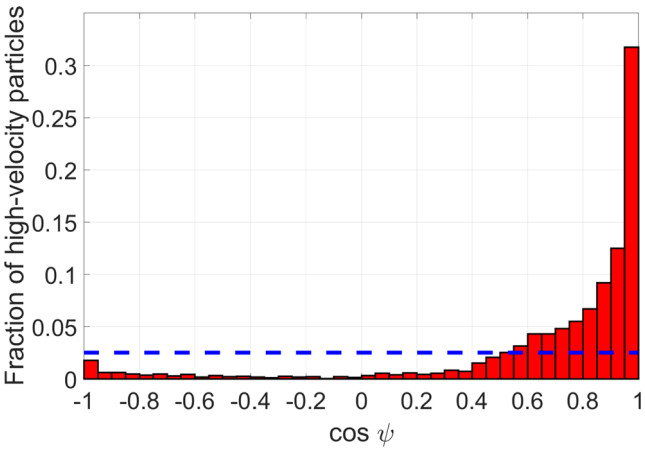


Figure 7. Histogram of $\cos\psi$ for all high-velocity test particles (above dashed orange line in Fig. 5) beyond 1.6 Mpc from the LG barycentre in order to best correspond to actual HVGs (Fig. 10). The dashed blue line indicates that 0.025 of the (weighted) particles should fall into each bin in $\cos\psi$ if their orbital poles were distributed isotropically.

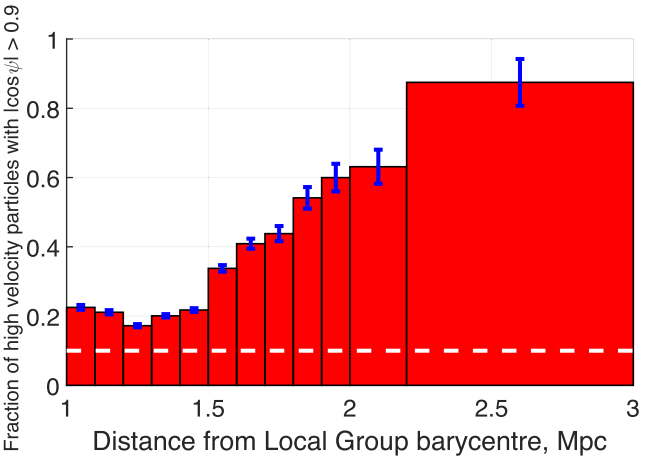


Figure 8. Fraction of high-velocity test particles (above dashed orange line in Fig. 5) in each radial bin which have $|\cos\psi| > 0.9$. The dashed white line shows the expectation for an isotropic distribution (0.1). Notice how the high-velocity particles furthest from the LG now tend to have a more anisotropic distribution. Uncertainties are estimated using binomial statistics, though this is not totally accurate due to our statistical weighting scheme (see the text). The result for the outermost radial bin is less reliable as we only have 24 particles in it, but other bins have at least 96 and should be quite reliable.

A pattern of this sort is apparent in the kinematics of the observed LG (Fig. 10). Despite appearances, our simulation indicates that only ~ 10 per cent of the particles at these distances belong to the upper branch, roughly consistent with the frequency of HVGs in Section 4.2. Other galaxies should lie below the Hubble flow due to the effect of gravity. In a MOND universe, they would probably lie closer to the Hubble flow than in our simulation due to the gravitational field of large-scale structures on the LG. This external field effect weakens the gravity exerted by the MW and M31 at long range (e.g. Banik & Zhao 2015). It is not caused by tides but arises because MOND gravity is non-linear in the matter distribution (equation 17).

In most parts of our simulated Hubble diagram, a moderate fraction of particles have orbits poorly aligned with the MW–M31 orbit (points coloured light blue, green or yellow in Fig. 5). However, this is not true for the high-velocity branch at distances $\gtrsim 1.5$ Mpc, which appears almost entirely dark red ($\cos\psi \approx 1$). This leads us to do a careful analysis of whether such particles really are distributed anisotropically.

Our initial grid of test particle positions is uniform in distance from the LG as well as in the spherical polar and azimuthal angles. Thus, each particle does not correspond to exactly the same volume/mass. We handle this by weighting each particle according to the volume it represents, which we find by integrating the usual spherical Jacobian factor over the range of initial coordinates covered by the particle[7].

We apply this weighting scheme to the high-velocity particles (above dashed orange line in Fig. 5) to determine their correctly weighted distribution over $\cos\psi$ (Fig. 7). If the particles have no preferred direction(s), then $\cos\psi$ should be distributed uniformly. Because we are investigating HVGs towards the edge of the LG, we also restrict to particles beyond 1.6 Mpc from its barycentre (as suggested by Fig. 10).

A large proportion of high-velocity particles appear to have very high values of $|\cos\psi|$. To see how robust this is, we determine the (weighted) fraction of particles in different radial bins with $|\cos\psi| > 0.9$, our proxy for an orbital plane almost aligned with that of the MW and M31. We estimate uncertainties using binomial statistics, which is only approximately correct here due to our weighting procedure. Apart from the outermost radial bin, there should be enough simulated particles in each one to accurately estimate this fraction.

Compared to an isotropic distribution, the fraction of nearly coplanar particles is very high (Fig. 8). This demonstrates that dwarfs flung out furthest from the LG should be distributed very anisotropically in a MOND context. Moreover, the preferred plane

[7] I.e. between the mid-points of the particle of interest and the particles at slightly smaller and larger polar angle, etc.

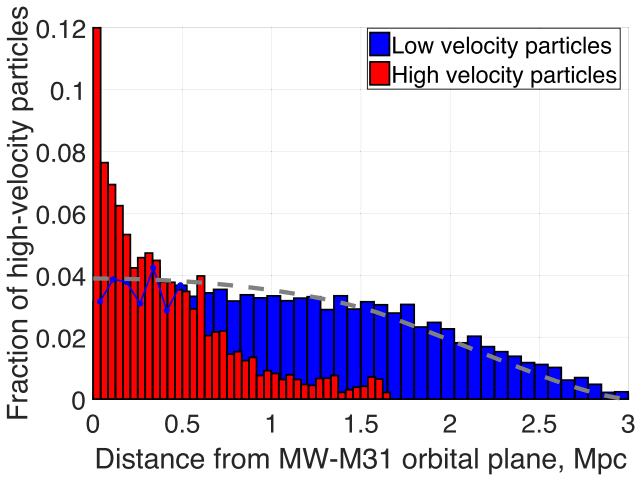


Figure 9. Histogram showing how far simulated particles are from the MW–M31 orbital plane. We only show particles currently at distances of 1.6–3 Mpc from the LG and sort them according to whether they are in the high-velocity branch of the Hubble diagram (above dashed orange line in Fig. 5). If they are, we show them as red. The remaining particles (shown in blue) are well described by an isotropic distribution (dashed grey line).

should correspond to the MW–M31 orbital plane. Within this plane, the HVGs should mostly be corotating with respect to the MW–M31 orbit (Fig. 7). However, it is not possible to test counter-rotation versus corotation at present due to a lack of accurate proper motions for the HVGs. This is why we focus on $|\cos\psi|$ rather than $\cos\psi$.

Gravitational slingshot interactions with the MW or M31 would be most efficient for particles flung out almost parallel to the motion of the perturber. Considering that the MW–M31 flyby occurred a fixed time in the past, these particles should currently be furthest away from the LG. Thus, it is not very surprising that the spatial distribution of such particles is highly flattened with respect to the MW–M31 orbital plane (Fig. 9).

Although this scenario almost exclusively leads to HVGs corotating with respect to the MW–M31 orbit, that is not always the case. For a particle flung out on an almost radial orbit with respect to the LG, only a small torque is needed to reverse the direction of its angular momentum. This may explain why the high-velocity test particles with $\cos\psi \approx -1$ tend to have rather small angular momenta. Pawlowski et al. (2011) suggested a similar mechanism to explain why some MW satellites like Sculptor are counter-rotating within the plane preferred by most remaining MW satellites (Piatek et al. 2006).

4 THE LOCAL GROUP IN ΛCDM

4.1 Refining the ΛCDM model

To better identify which galaxies may have been flung out by a fast-moving MW/M31 in the way discussed in Section 3, we refine our previous ΛCDM dynamical model of the LG (Banik & Zhao 2017). In this work, we use an updated input catalogue. The main changes are a more accurate distance measurement to NGC 404 (Dalcanton et al. 2009) and to Leo P (McQuinn et al. 2015). For NGC 4163, we use a less accurate distance of 2.95 ± 0.07 Mpc to bracket the range between the measurements of Dalcanton et al. (2009) and Jacobs et al. (2009), both of which are based on data from the *Hubble Space Telescope*.

We also make improvements to the procedure used to find the best-fitting model parameters. The flatline level of the MW rotation curve is no longer assumed equal to its amplitude $v_{c,\odot}$ at the position of the Sun.[8] We let the former vary with a prior of $205 \pm 10\,\mathrm{km\,s^{-1}}$ (McGaugh 2016a), while $v_{c,\odot}$ is fixed at $232.8\,\mathrm{km\,s^{-1}}$ (McMillan 2017). The time resolution is improved $10\times$ so that 5000 steps are now used to cover the history of the Universe since redshift 9 ($a = 0.1$), leading to a much better handling of close encounters.

The best-fitting solution is found by applying gradient descent to all model parameters, using a method similar to that described in Section 5.1. To maximize the chance of matching observations, we run a grid search through the trajectories of all the dwarf galaxies, which are treated as test particles. Because the algorithm uses a least action method (Peebles et al. 2011), trajectories are solved by relaxing an initial guess towards a solution that satisfies the equations of motion. The initial guess has the comoving position varying linearly with a. The direction and magnitude of the present peculiar velocity of each dwarf are varied over a 3D grid of possibilities, giving the algorithm a much better chance of finding slingshot encounters that might otherwise get missed if the initial trajectory went nowhere near the space–time location of the encounter. The issue of local but not global minima can always be solved with a grid search, which in this case is feasible for the dwarf galaxies because the trajectory of each one does not influence the gravitational field in the LG and thus the trajectory of anything else.[9]

As some improvements are indeed found in this way, we repeat the gradient descent stage and the grid search in an alternating manner until the algorithm converges in the sense that the grid search stops improving the agreement between model and observations. This process takes a few days and yields reliable trajectories for all simulated galaxies – their present-day positions and velocities are almost perfectly recovered (maximum errors of 9 pc and 16 m $\mathrm{s^{-1}}$, respectively) if we solve their trajectories forwards using a more traditional fourth-order Runge–Kutta method with $10\times$ finer resolution.

Using equation (30) from Banik & Zhao (2017), we adjust the predictions of this best-fitting model for the effect of tides raised on the LG by the GA. This only slightly affects our results, which are shown in Fig. 10. The main difference from our previously published results (Banik & Zhao 2017, fig. 13) is that Tucana is now consistent with ΛCDM expectations.

At distances $\gtrsim 1.5$ Mpc, a bimodal distribution of ΔGRVs is apparent, similar to that in our MOND simulation of the LG (Fig. 5). Moreover, the galaxies in the lower branch predominantly have $\Delta\mathrm{GRV} < 0$, perhaps a sign of the stronger gravity in MOND than in the ΛCDM model whose predictions have been subtracted.

4.2 Selecting high-velocity galaxies

To find HVGs in the real LG, we compare the distances d of our target galaxies from the LG barycentre with their $\Delta\mathrm{GRV} \equiv \mathrm{GRV_{obs}} - \mathrm{GRV_{model}}$ relative to our best-fitting 3D dynamical model of the LG (Fig. 10). We expect that these dwarf galaxies were flung out at high speed by the MW or M31, implying they passed close to the space–time location of the MW–M31 flyby. Thus, such dwarfs

[8] This is sometimes called the local standard of rest.
[9] We did not do a full grid search as that would involve jointly varying trajectories of all 21 massive particles, impossible in a reasonable time frame.

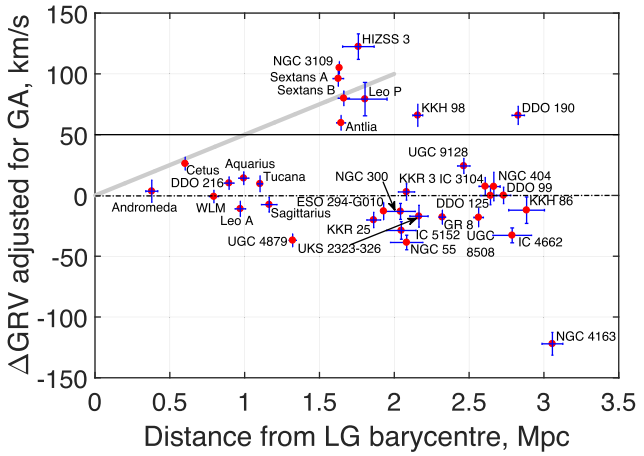


Figure 10. The deviation ΔGRV of each target galaxy from our best-fitting ΛCDM model is shown against its distance from the LG barycentre. If the model worked perfectly, then all galaxies would have ΔGRV ≡ 0 (dot–dashed line) as model predictions are subtracted. Given likely model uncertainties of ∼25 km s^{-1} (Aragon-Calvo et al. 2011), ΛCDM would thus find it difficult to explain galaxies with ΔGRV > 50 km s^{-1} (above solid black grid line). In our MOND scenario of a past MW–M31 flyby, we expect the HVGs to broadly follow a trend of 50 km s^{-1} Mpc^{-1} (diagonal grey line) and to reach distances up to ∼2 Mpc (Fig. 5).

should follow a ΔGRV ≡ d relation of the sort apparent in our MOND simulation of the LG (Fig. 5). A relation like this is evident in Fig. 10, where we have added a solid grey line at $u = 50$ km s^{-1} Mpc^{-1} to make it clearer. An RV excess of this magnitude suggests that the MW–M31 flyby occurred $\sim (H_0 + u)^{-1} \approx 8$ Gyr ago. This is consistent with their expected orbital evolution in MOND (Fig. 4).

A larger MW–M31 pericentre would not affect this conclusion much as a HVG would still require a similar velocity to reach its presently observed position from the space–time location of the MW–M31 flyby, the timing of which is constrained observationally if we assume this event led to the formation of the MW thick disc (Quillen & Garnett 2001). However, a weaker MW–M31 encounter would reduce the maximum distance at which we might expect to see a HVG.

As our ΛCDM-based model of the LG is not a perfect representation of a ΛCDM universe, we expect model uncertainties of ∼25 km s^{-1} based on how an LG analogue deviates from spherical symmetry in a detailed cosmological simulation of ΛCDM (Aragon-Calvo et al. 2011).[10] Thus, we focus our attention on the galaxies with ΔGRV > 50 km s^{-1} and following the ΔGRV ∝ d relation. This leads to the HVG sample in Table 3.

The reasonably high ΔGRV of DDO 190 (66 ± 7 km s^{-1}) is still marginally compatible with our model if we assume a model uncertainty of ∼25 km s^{-1}. This is especially true when considering that the much larger distance of DDO 190 from the LG barycentre suggests that it should have a much higher ΔGRV if it really was flung out in the same way as e.g. NGC 3109. Thus, we do not consider DDO 190 as being a genuine HVG, even though its ΔGRV is slightly on the high side.

Although it would be quite normal to have one such instance amongst our 34 LG target galaxies of observations exceeding model predictions by 2.6σ (probability ≈ 0.13), a second such instance would be unexpected. Thus, it would be rare to observe ΔGRVs as large as for DDO 190 and KKH 98 (66 ± 9 km s^{-1}) if we treat both as having normal kinematics in a ΛCDM context. Given that DDO 190 deviates very substantially from the ΔGRV ∝ d relation typically followed by HVGs, this suggests that KKH 98 may be an HVG. Although it does not fit the ΔGRV ∝ d relation perfectly, some scatter about this is expected because the LG is not spherically symmetric and is presently observed from an off-centre vantage point. Because most HVGs lie at rather similar angles to the MW–M31 line, we expect larger deviations from this relation for HVGs like KKH 98 which lie at a totally different angle (Fig. 16).

For the particular case of KKH 98, its position rather close to the MW–M31 line means that its ΔGRV would be more sensitive to where our model puts the centre of mass of the LG i.e. its preferred MW:M31 mass ratio. If the MW and M31 masses are not equal but only 0.3 of their total mass is in the MW (Section 3), then the LG barycentre would be shifted by ∼160 kpc towards M31 and thus by a similar amount towards KKH 98. This would put it closer to the LG barycentre. In the LG, the RV rises with distance at a rate close to 100 km s^{-1} Mpc^{-1} due to the gravity of the MW and M31 (Banik & Zhao 2017, fig. 5). Thus, a galaxy 160 kpc closer to the LG barycentre should be receding away from it 16 km s^{-1} slower. Moreover, the MW would be moving 22 km s^{-1} faster towards M31 (and thus KKH 98) given the observed MW–M31 relative RV (van der Marel et al. 2012a). A more massive M31 would also be expected to reduce GRVs of objects in the general vicinity of KKH 98 compared to a situation where the MW and M31 have equal mass. Even without this dynamical effect, the kinematic effects alone would reduce the predicted GRV of KKH 98 by ∼40 km s^{-1}, but would have a smaller effect on the other HVGs and DDO 190 as their sky positions are almost orthogonal to the MW–M31 line (Fig. 16). If this is correct, it explains why the ΔGRV of KKH 98 falls below the ΔGRV ∝ d relation by about this much.

Our scenario implies that any plane passing close to most of the HVGs should also pass close to the MW and M31. Thus, we apply our plane-fitting algorithm (Section 5.1) to the galaxies listed in Table 3, always including the MW and M31 in our sample. We exclude HIZSS 3 despite the fact that it should be treated as an HVG because a much thinner plane is obtained without it (Fig. 11). Naturally, it would be more common to find a sample of galaxies with an anisotropic distribution if one of them can be removed arbitrarily with the explicit intention of making the remaining ones have a more anisotropic distribution. We account for this using the method in Section 5.2. There are also good observational reasons for excluding HIZSS 3 from our analysis, in which case this 'look elsewhere' effect should not be considered (Section 6.3).

Table 3. Galaxies considered when finding the plane best fitting the high ΔGRV galaxies in our sample, which we select based on Fig. 10.

Galaxies included in our plane fit	Distance from MW–M31 mid-point (Mpc)	ΔGRV (km s^{-1})
Milky Way	0.382 ± 0.04	NA
Andromeda	0.382 ± 0.04	3.5 ± 9.1
Sextans A	1.624 ± 0.036	96.1 ± 6.3
Sextans B	1.661 ± 0.037	79.9 ± 6.0
NGC 3109	1.631 ± 0.014	105.0 ± 5.3
Antlia	1.642 ± 0.030	59.7 ± 6.1
Leo P	1.80 ± 0.15	79 ± 14
KKH 98	2.160 ± 0.033	65.5 ± 9.1

[10] This is discussed in more detail in section 4 of Banik & Zhao (2016).

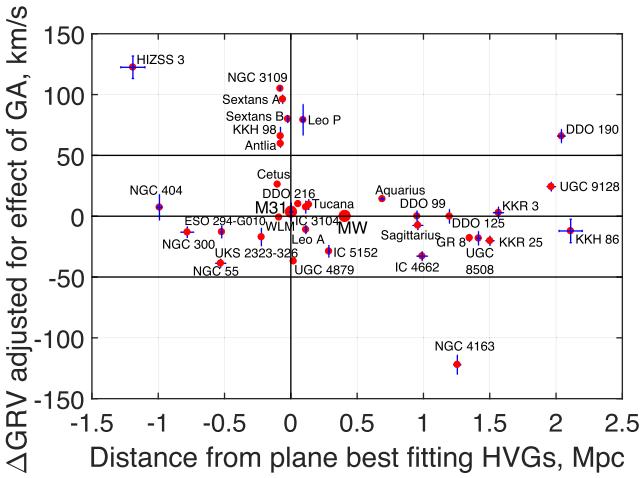


Figure 11. ΔGRVs of target galaxies are shown against their offsets from the best-fitting plane through the ones with the largest ΔGRVs except HIZSS 3 (parameters of this plane given in central column of Table 4). By definition, ΛCDM predicts ΔGRV = 0 (central horizontal line) with an uncertainty of $\sim$25 km s^{-1} (Aragon-Calvo et al. 2011). Thus, the model cannot easily explain galaxies with ΔGRV $>$ 50 km s^{-1} (above upper horizontal line). Most of these galaxies lie very close to a plane (near vertical grid line), unlike the rest of our sample. For the MW, the concept of a ΔGRV is meaningless, so we show this as 0.

5 ANALYSING THE LOCAL GROUP

5.1 Finding the best-fitting plane

To quantify whether a set of galaxies is distributed anisotropically, we need to define a measure of anisotropy and determine how unusual its value is. The statistic we will use is $z_{\rm rms}$, the root mean square (rms) of the minimum distances between the galaxies we consider and the best-fitting plane through them (i.e. the one that minimizes $z_{\rm rms}$). With respect to a plane having normal $\hat{\boldsymbol{n}}$ and containing the vector $\boldsymbol{r}_0$, the vertical dispersion is

$$z_{\rm rms}{}^2 = \frac{1}{N}\sum_{i=1}^{N}\left[(\boldsymbol{r}_{\rm i}-\boldsymbol{r}_0)\cdot\hat{\boldsymbol{n}}\right]^2 \tag{31}$$

$$= \hat{\boldsymbol{n}}\cdot\left(\boldsymbol{I}\hat{\boldsymbol{n}}\right) \text{ where} \tag{32}$$

$$\boldsymbol{I}_{\rm jk} \equiv \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{r}_{\rm i}-\boldsymbol{r}_0)_j\,(\boldsymbol{r}_{\rm i}-\boldsymbol{r}_0)_{\rm k} \tag{33}$$

The galaxies are at heliocentric positions $\boldsymbol{r}_{\rm i}$. The minimum of $z_{\rm rms}$ is attained when $\boldsymbol{r}_0 = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{r}_{\rm i}$, corresponding to the geometric centre of the N galaxies to which we are trying to fit a plane. We find the best-fitting orientation $\hat{\boldsymbol{n}}$ using a gradient descent method (e.g. Fletcher & Powell 1963). The gradient of $z_{\rm rms}{}^2$ with respect to its argument $\hat{\boldsymbol{n}}$ is $\frac{2}{N}\left(\boldsymbol{I}\hat{\boldsymbol{n}}\right)$ less the component of this parallel to $\hat{\boldsymbol{n}}$. At the minimum of $z_{\rm rms}$, its gradient vanishes, implying that $\hat{\boldsymbol{n}}$ is an eigenvector of the inertia tensor $\boldsymbol{I}$ corresponding to its minimum eigenvalue. This provides a non-iterative way of minimizing $z_{\rm rms}$, taking advantage of the characteristic polynomial of $\boldsymbol{I}$ being a cubic whose roots can be found analytically. However, we find that this approach is slower than gradient descent, a much more general method which we also used in Section 4.1.

We minimize issues of local minima by starting the gradient descent based on whichever $\hat{\boldsymbol{n}}$ yields the smallest $z_{\rm rms}$ in a low-resolution grid of possible directions for $\hat{\boldsymbol{n}}$. Once the angular step size is below 0.°006, we stop doing further iterations. As well as ensuring that our algorithm always converges in this sense, we also verify it using mock data designed to lie close to a plane with known $\hat{\boldsymbol{n}}$ and $z_{\rm rms}$. We are always able to accurately recover their input values.

5.2 Statistical analysis

The MW, M31 and all but one of the HVGs lie close to a plane (Fig. 11). We need to reflect this when determining the likelihood of $z_{\rm rms}$ being as low as the observed 101 kpc. To see if this is consistent with isotropy, we conduct a series of Monte Carlo (MC) trials in which we randomize the directions to these galaxies and recompute $z_{\rm rms}$. Thus, the probability distribution of the Galactic longitude l is uniform while that of the Galactic latitude b is

$$P(b)\mathrm{d}b = \frac{1}{2}\cos b\mathrm{d}b \tag{34}$$

To mimic uncertainties in measured distances to LG galaxies, we randomly vary their heliocentric distances using Gaussian distributions of the corresponding widths. In the very rare cases where this yields a negative distance, we set this to 0.

To account for HIZSS 3 being excluded from our plane fit despite its high ΔGRV, we use the procedure described in Section 5.1 to find the best-fitting plane through every combination of all HVGs but one as well as the MW and M31.[11] The combination yielding the lowest $z_{\rm rms}$ is considered the analogue of the observed HVG system less HIZSS 3 for that particular randomly generated mock catalogue. In Section 6.5, we perform calculations where we select the combination yielding the lowest aspect ratio A rather than $z_{\rm rms}$.

$$A \equiv \frac{z_{\rm rms}}{\sqrt{{r_{\rm rms}}^2 - {z_{\rm rms}}^2}} \text{ where} \tag{35}$$

$${r_{\rm rms}}^2 \equiv \frac{1}{N}\sum_{i=1}^{N}|\boldsymbol{r}_{\rm i}-\boldsymbol{r}_0|^2 = \mathrm{Trace}\,(\boldsymbol{I}) \tag{36}$$

$r_{\rm rms}$ is the rms distance of the galaxies from their geometric centre $\boldsymbol{r}_0$. To get the rms extent of the system after projection into the best-fitting plane, we need to subtract $z_{\rm rms}$ in quadrature. Dividing $z_{\rm rms}$ by the result then gives a measure of the typical 'vertical' extent of galaxies out of this plane relative to their 'horizontal' extent within it. We would obtain identical probabilities for the observed situation if we defined A as $\frac{z_{\rm rms}}{r_{\rm rms}}$ instead, as long as it is defined in the same way for the actual HVGs and the mock sample in each MC trial (Section 5). This is because A is a monotonic function of $\frac{z_{\rm rms}}{r_{\rm rms}}$ with either definition.

The major LG galaxies along with the HVGs except HIZSS 3 (full list in Table 3) define a rather thin plane whose parameters are given in Table 4. This allows us to compare the ΔGRV of each galaxy[12] with its minimum distance from this plane. The galaxies in our full sample have a wide range of positions relative to it, with a similar number on either side (Fig. 11). However, the HVGs tend to lie very close to it. The only exception is HIZSS 3, justifying our decision not to consider it when defining the HVG plane. In any case, the observations for HIZSS 3 are rather insecure (Section 6.3).

In Table 5, we give the criteria which we use to determine whether the distribution of HVGs in an MC trial is analogous to their observed distribution. We choose these criteria so that they should be satisfied if the LG behaves similarly to our MOND simulation of it

[11] It would not make sense for the major LG galaxies to lie far from this plane as it should be their mutual orbital plane.
[12] Adjusted for the GA using equation (30) of Banik & Zhao (2017).

Table 4. Information about the plane best fitting the galaxies listed in Table 3, with distances in kpc and plane normal direction in Galactic coordinates (latitude last). The last column shows how our results change if Antlia is removed from our sample as it could be a satellite of NGC 3109 (van den Bergh 1999). The effect on our statistical analysis is described in Section 6.4.

Quantity	Full sample	Without Antlia
Galaxies in plane	8	7
Normal to plane of high ΔGRV galaxies	$\begin{bmatrix} 204.^{\circ}4 \\ -30.^{\circ}1 \end{bmatrix}$	$\begin{bmatrix} 206.^{\circ}6 \\ -31.^{\circ}8 \end{bmatrix}$
rms plane width (kpc)	101.1	101.9
Aspect ratio (equation 35)	0.0763	0.0750
MW offset from plane	224.7	195.4
M31 offset from plane	−0.6	−12.8
Angle of MW–M31 line with plane	16.°2	14.°9

Table 5. Criteria used to judge whether a randomly generated population of galaxies is analogous to the observed population of HVGs. Only one of the first two criteria is used at a time. Note that the criteria are not all independent. For example, as the MW and M31 positions are fixed and our plane fitting procedure always considers them, the thinnest planes are likely to be obtained when these galaxies are close to the plane best fitting the HVGs. This makes it more likely that the 'barycentre offset' criterion is satisfied (see corners of Table 6).

Criterion	Meaning
Plane thickness	There must be a plane of HVGs with rms thickness (equation 31) below that observed
Aspect ratio	There must be a plane of HVGs with aspect ratio (equation 35) below that observed
Barycentre offset	Barycentre of MW and M31 (assuming 30 per cent of total mass in MW) closer to plane than observed situation
Direction	Normal to HVG plane closer to expected direction (Section 2.2) than observed

(Section 3). Depending on which measure of anisotropy is used, we consider one of the first two criteria alongside both of the others.

In Section 2.2, we used a toy model to find the MW–M31 orbital pole $\widehat{\boldsymbol{h}}$ leading to a past close encounter between these galaxies in the most favourable orientation for the formation of their satellite planes. One might expect the HVGs to define this orbital plane. Thus, it is interesting that there is only a $\sim 19^{\circ}$ angle between the plane normal defined by the HVGs and our expected direction for $\widehat{\boldsymbol{h}}$.

Another requirement of our model is that the HVG plane should intersect the MW–M31 barycentre, which we take to be 0.3 of the way from M31 towards the MW for reasons discussed at the start of Section 3.1.[13] This puts the MW–M31 barycentre 67 kpc from the best-fitting plane, a rather small offset from a plane with a radial extent of ∼1.3 Mpc.

6 RESULTS

Applying the criteria defined in Table 5 to 20 million MC trials based on our nominal sample of HVGs (Table 3), we obtain the results shown in Table 6. The uncertainties are found by repeating the MC trial using four different seeds for the random number generator, with each seed used for 5×10^6 trials.[14] The variance between the results gives an indication of the uncertainty in our final result, which is a simple mean. We also estimate the error using binomial statistics. Our final error estimate is based on whichever method suggests a higher uncertainty. Usually, this is the method involving comparing results obtained using different random number seeds. In all cases, we are able to constrain the proportion of MC trials matching our criteria to within a few per cent of its most likely value.

Table 6. MC trial-based probabilities in parts per thousand (‰) of the HVGs (all galaxies in Table 3) matching various combinations of the criteria defined in Table 5. These criteria are used to determine if a mock system of HVGs is analogous to the observed system, using the method outlined in Section 5.2. When the same criterion appears in both the row and column headings, the result is the probability of matching that criterion alone, regardless of the others. The probability of all three criteria being met simultaneously is 1.48 ± 0.10‰, which corresponds to the first row of Table 7.

	Thickness	Direction	Barycentre offset
Thickness	4.6 ± 0.3		
Direction	2.3 ± 0.1	417.4 ± 0.5	
Barycentre offset	2.4 ± 0.2	81.1 ± 1.3	181.8 ± 2.5

The direction criterion proved the least problematic due to the rather wide range of allowed orientations of the plane best fitting the mock galaxies. This criterion was met ∼417‰ of the time.

The plane of HVGs is offset from the MW–M31 barycentre by 67 kpc, which is rather small considering the extent of the HVG plane (∼1 Mpc). Thus, the 'barycentre offset' criterion in Table 5 is only met around 182‰ of the time.

By far the most important criterion is the requirement that all but one mock HVG define a plane with rms thickness smaller than observed. This criterion is met only 5.2 ± 0.2‰ of the time. Consequently, it is very unlikely (probability 1.48 ± 0.10‰) that all three criteria are satisfied simultaneously.

To check how various assumptions affect our results, we conduct several versions of our statistical analysis. We describe these variations next and summarize our findings in Table 7.

6.1 Altering the expected MW–M31 orbital plane

In Section 2.2, we used a toy model for the origin of the MW and M31 satellite planes to estimate $\widehat{\boldsymbol{h}}$, the direction of the MW–M31 orbital angular momentum. We parametrized this using the angle θ, which we take to be 70°. The actual value may well be different, making it important to know if this affects our conclusions. To this end, we repeat the analysis shown in Table 6 with $\theta = 75^{\circ}$. A slightly different expected $\widehat{\boldsymbol{h}}$ alters the range of 'allowed' orientations for the best-fitting plane. However, the results hardly change (Table 7).

This is because a fairly wide range of plane normal directions are allowed such that this is not the main difficulty with matching the observed situation (Table 6). Instead, the difficulty lies in obtaining a plane as thin as the observed 101 kpc despite the HVGs having much larger distances (Table 3). Thus, in this contribution, we use

[13] Even in ΛCDM, it is likely that M31 has a higher mass than the MW (Peñarrubia et al. 2014; Banik & Zhao 2016).

[14] We use the rng('shuffle') command in MATLAB to initialize its random number generator based on the date and time. We verify that different runs give slightly different results, which is inevitable for runs started at least a few seconds apart.

Table 7. How our results depend on various model assumptions. The final column shows the probability of an MC trial satisfying the criteria given in Table 5 based on randomizing the directions towards the HVGs in Table 3. Galaxies whose names have been crossed out are excluded from our sample in that particular investigation, with the nominal sample corresponding to the 'full sample' column of Table 4. Unless stated otherwise, we use the 'plane thickness' criterion and a MW–M31 orbital plane corresponding to $\theta = 70^\circ$ in Fig. 3, with 0.3 of the total MW and M31 mass assumed to be in the MW (Section 3.1).

Investigation	Sample	Probability (‰)
Nominal (physical thickness)	All	1.48 ± 0.10
$\widehat{\boldsymbol{h}}$ rotated 5° south ($\theta = 75^\circ$)	All	1.51 ± 0.10
Distances fixed	All	1.45 ± 0.01
Nominal	~~HIZSS 3~~	0.41 ± 0.02
Nominal	~~Antlia~~	5.17 ± 0.36
Aspect ratio	All	1.62 ± 0.01
Aspect ratio	~~Antlia~~	5.35 ± 0.02

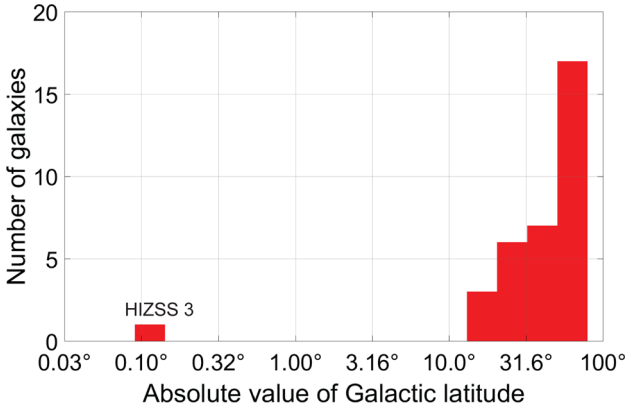


Figure 12. Histogram showing the Galactic latitudes of all galaxies in our sample (not just HVGs). Notice the very small value for HIZSS 3 (0.°09, Massey et al. 2003).

$\theta = 70^\circ$ as we consider this to be more realistic. Although this does not much affect our results, a lower value for θ leads to a greater torque on the M31 disc, which is otherwise quite small. This is evident from our toy model preferring a solution close to the gap in Fig. 3 arising from the MW–M31 line at their closest approach lying almost within the M31 disc plane (equation 9). Thus, we think it will be easier for a MOND simulation to reproduce the observed properties of the MW and M31 discs and satellite planes using a slightly smaller value of θ or equivalently with $\widehat{\boldsymbol{h}}$ pointing slightly closer to the Galactic Equator.

6.2 Using fixed distances

Our analysis randomizes both the directions towards the HVGs as well as their distances. The latter are drawn from a Gaussian distribution corresponding to the relevant measurement and its uncertainty, with any negative mock distances converted to 0. In the same way, we also vary the heliocentric distances to M31 (783 ± 25 kpc, McConnachie 2012) and to the centre of our own Galaxy (8.20 ± 0.09 kpc, McMillan 2017). To see how this affects our results, we repeat our analysis using fixed distances to all galaxies in our sample. This has only a minor impact on our results, with the probability of matching all criteria falling from 1.48 ± 0.10‰ to 1.45 ± 0.01‰. This is consistent with no change, suggesting that distances to LG galaxies are known accurately enough for our analysis.

6.3 Excluding low galactic latitudes

We do not expect that *all* HVGs necessarily lie close to a plane (Fig. 9). Thus, it is not too surprising that one of them (HIZSS 3) lies far outside the plane defined by the others (Fig. 11). However, it is also possible that there are observational issues for HIZSS 3 due to its extremely low Galactic latitude ($b = 0.09^\circ$, Massey, Henning & Kraan-Korteweg 2003). It is readily apparent that no other galaxy in our entire sample (not just the HVGs) have sky positions closer to the Galactic disc (Fig. 12).

The low Galactic latitude of HIZSS 3 may reduce the accuracy of its distance and/or RV measurement. In particular, its tip of the red giant branch-based distance (Silva et al. 2005) seems rather insecure as this feature on its colour–magnitude diagram (CMD) is not very well defined (see their fig. 6). It is based on only a small number of stars in images suffering from a rather high contamination fraction by foreground MW stars due to the low $|b|$.

The authors also mention possible complications in the dust correction due to only reliably knowing this for the central part of HIZSS 3 but needing to use imaging over a much larger area of it in order to get enough stars in its CMD. In particular, the distance estimate would be higher if its central regions were particularly dusty, whereas a uniform dust correction was assumed. A more distant galaxy would have a lower ΔGRV. This could be clarified with an updated distance to HIZSS 3, but unfortunately we only have one measurement that is now more than a decade old.

As well as contamination by foreground stars, Galactic dust is a major issue at very low Galactic latitudes. Due to the large amount of dust along such lines of sight, it becomes crucial to correct for this accurately. The foreground extinction measure $E(B - V)$ was estimated at 1.32 by Silva et al. (2005). However, more recent work suggests an extinction of only 0.88 (Schlafly & Finkbeiner 2011), with an uncertainty close to 0.1 in both cases. Reducing the extinction towards an astrophysical object increases its apparent luminosity. For consistency with observations, it must be further away than previously thought, in this case by ~0.4 Mpc. This would increase the predicted GRV of HIZSS 3 by a substantial amount. A look at fig. 5 of Banik & Zhao (2017) suggests that the increase would be ~50–100 km s^{-1}, perhaps resolving the discrepancy between the observed GRV of HIZSS 3 and that expected in our model. If this is correct, then HIZSS 3 is not a HVG.

Silva et al. (2005) often referred to a basic (1D) dynamical model of the LG combined with the redshift of HIZSS 3 to try and justify their distance estimate. However, the law of gravity governing the LG and its past dynamical history are presently unclear, in particular whether there was a past close MW–M31 encounter (Zhao et al. 2013). This makes it important to obtain redshift-independent distances. If this is not possible, then we note that HIZSS 3 is likely not a HVG as only a minority of galaxies are (e.g. if we use a threshold of $\Delta\text{GRV} = 50$ km s^{-1} in Fig. 10). This is possible if it is further away than we assumed, as seems likely.

Apart from HIZSS 3, the galaxy in our sample with the lowest $|b|$ is the Sagittarius dwarf irregular galaxy ($b = -16.°3$, Longmore et al. 1978). This suggests that observations at lower $|b|$ are difficult. Thus, we repeat our analysis with the directions towards the HVGs randomized but restricted such that $|b| \geq 15^\circ$. We consider this a plausible range of 'unobservable' Galactic latitudes as a limit of 19.°47 was used by Cautun et al. (2015) when dealing with MW

satellites, though López-Corredoira & Kroupa (2016) considered this a bit too high.

The major consequence of such a restriction is that HIZSS 3 must be removed from our sample. We think this is reasonable given the above reasons for why its distance measurement seems rather unreliable. However, no other galaxy is removed from our sample (Fig. 12). This leads to *all* HVGs lying very close to a plane, greatly reducing the probability of obtaining a situation matching the observed one (probability decreases from 1.48 ± 0.10‰ to 0.41 ± 0.02‰). This is mainly because the chance of getting a plane as thin as observed drops from 4.6 ± 0.3‰ to 0.85 ± 0.03‰.

6.4 Excluding Antlia

We have treated all HVGs as independent LG dwarf galaxies. In particular, we assumed that NGC 3109 and Antlia are unrelated objects. However, they may be gravitationally bound (van den Bergh 1999). There are indications that they have recently interacted, based on observations of both NGC 3109 (Barnes & de Blok 2001) and Antlia (Penny et al. 2012). This is more likely if Antlia is a satellite of NGC 3109. The $41 \pm 1\,\mathrm{km\,s^{-1}}$ difference in their RVs (Karachentsev, Makarov & Kaisina 2013) and their 1°.19 sky separation (corresponding to $\gtrsim 28$ kpc) seem consistent with this scenario given that the distance to Antlia (1.31 ± 0.03 Mpc, Pimbblet & Couch 2012) is almost the same as that to NGC 3109 (1.286 ± 0.015 Mpc, Dalcanton et al. 2009). Thus, the galaxies are probably $\lesssim 40$ kpc apart and may well be bound. Indeed, the latest results indicate that this is rather likely (Kourkchi & Tully 2017, section 9).

To account for this possibility, we exclude Antlia from our sample as it almost certainly has a much lower mass than NGC 3109 given that it is ∼5 mag fainter (McConnachie 2012, table 3). Although this hardly alters the properties of the plane best fitting the remaining HVGs (Table 4), the loss of a HVG increases the probability that the remaining ones end up close to a plane in a random MC trial. As a result, the proportion of MC trials meeting all our criteria more than triples but is still only 5.2 ± 0.4‰. Even this rather small figure assumes that HIZSS 3 has an accurate distance measurement and so should be in our sample of HVGs. As discussed in Section 6.3, this may well be incorrect, further reducing the proportion of MC trials matching observations.

6.5 Aspect ratio instead of physical thickness

So far, our results have been dominated by how likely it is to obtain a plane of HVGs with a physical thickness (equation 31) as small as observed. Another measure of anisotropy is the aspect ratio A (equation 35). To see how sensitive our results are to which statistic is used, we repeat our nominal analysis using A instead of $z_{\rm rms}$, each time considering the combination of all but one HVG that yields a best-fitting plane with the smallest A. This very slightly increases the proportion of MC trials matching all three criteria, from 1.48 ± 0.10‰ to 1.62 ± 0.01‰. This is consistent with no change, suggesting that our results are not much dependent on which statistic is used to quantify the anisotropy of the HVGs.

6.6 Using a flattened prior distribution

Our results indicate that the HVGs are likely not distributed isotropically (Table 7). However, ΛCDM typically produces filamentary structures and sheets (e.g. Noh & Lee 2007). This leads us to investigate whether a slightly flattened distribution of HVGs would

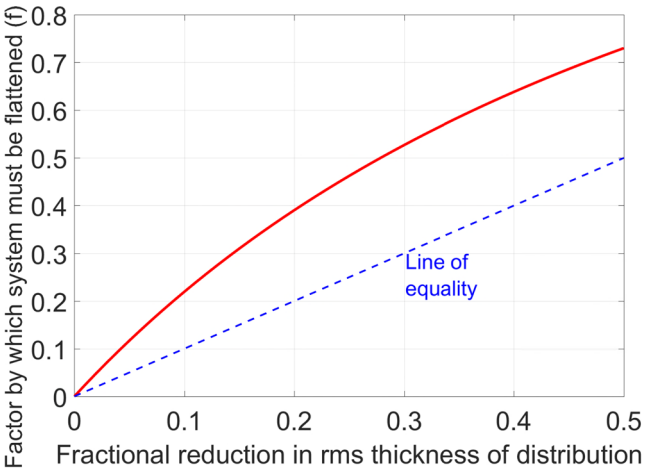


Figure 13. The solid red line shows the factor by which mock galaxy positions must be 'flattened' (f in equation 37) as a function of the desired fractional reduction in the rms thickness of an initially isotropic distribution of particles kept at fixed distances. This last requirement implies a non-trivial relation between the quantities plotted as any $f > 0$ generally implies a reduction in distance to an object, requiring a rescaling of its position vector (equation 38). As a result, the fractional reduction in rms thickness is much smaller than f. This is evident from our results lying left of the line of equality (dashed blue).

be consistent with observations. For this purpose, we randomly[15] select a vector $\widehat{\boldsymbol{n}}$ and 'flatten' the heliocentric position vectors $\boldsymbol{r}$ towards each mock HVG as well as to M31 and the centre of the MW using

$$\widetilde{\boldsymbol{r}} = \boldsymbol{r} - f\left(\boldsymbol{r}\cdot\widehat{\boldsymbol{n}}\right)\widehat{\boldsymbol{n}} \tag{37}$$

A straightforward application of equation (37) would put these objects at a smaller distance than observed. This is resolved by a subsequent rescaling of the position vector.

$$\widetilde{\boldsymbol{r}} \rightarrow \frac{|\boldsymbol{r}|}{|\widetilde{\boldsymbol{r}}|}\widetilde{\boldsymbol{r}} \tag{38}$$

This radial rescaling means that the final value of $z_{\rm rms}$ is reduced by a smaller fraction than f. Taking an isotropically distributed set of unit vectors $\widehat{\boldsymbol{r}}$, the variance in the component of $\widetilde{\boldsymbol{r}}$ along $\widehat{\boldsymbol{n}}$ is

$${z_{\rm rms}}^2 = \int_0^{\frac{\pi}{2}} \frac{\alpha^2\cos^2\theta}{\alpha^2\cos^2\theta + \sin^2\theta}\sin\theta\,\mathrm{d}\theta \tag{39}$$

$$= \frac{\alpha^2}{\beta^2}\left[\frac{1}{2\beta}\ln\left(\frac{1+\beta}{1-\beta}\right) - 1\right] \text{ where} \tag{40}$$

$$\alpha \equiv 1 - f \text{ and} \tag{41}$$

$$\beta \equiv \sqrt{1-\alpha^2} \tag{42}$$

The integral can be solved by substituting for $\beta\cos\theta$. This lets us determine the value of f required to obtain a desired fractional reduction in the $z_{\rm rms}$ of an isotropically distributed unit vector via the successive application of equations (37) and (38). We achieve this using a Newton–Raphson root finder. Our results are shown in Fig. 13.

In Fig. 14, we show the results of repeating our MC analysis using flattened prior distributions for the HVGs. We cover priors ranging from spherical to moderately oblate, similar to the range expected

[15] From an isotropic distribution.

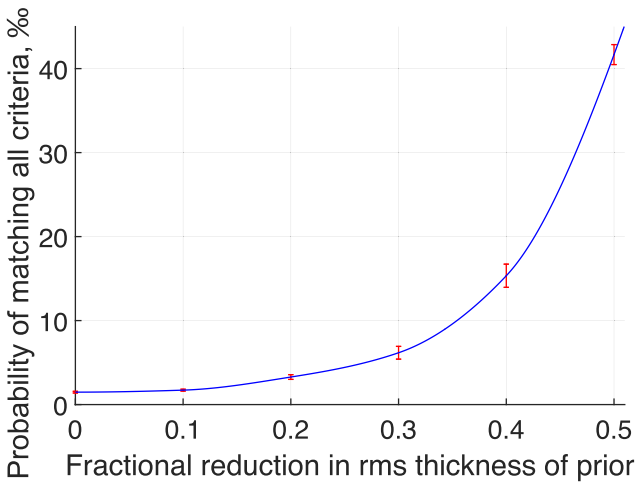


Figure 14. The effect on our MC analysis of flattening our prior for the spatial distribution of mock galaxies. The value of f used in equation (37) can be found using Fig. 13. The probabilities shown are those of obtaining a mock catalogue of HVGs with properties analogous to the observed ones, according to the criteria defined in Table 5.

in ΛCDM for individual haloes (Butsky et al. 2016). Considering the positions of the galaxies in our full sample (not just HVGs), this seems a reasonable hypothesis for the LG (Fig. 11). However, a mildly flattened underlying distribution is inconsistent with the observed HVG system as that has an aspect ratio <0.1 (Table 4). Our results suggest that a flattening of $\gtrsim$0.5 would be required for consistency with observations.

To test whether a flattening of this sort is likely to be present in the LG, we apply our plane-fitting procedure (Section 5.1) to our complete sample of LG galaxies, not just the HVGs. We find a preference for some flattening with respect to the axis (210°, −6°) in Galactic coordinates, though the rms thickness of the best-fitting plane is rather high at 645 kpc. The aspect ratio of our whole sample is 0.358, close to half the value of $\frac{1}{\sqrt{2}}$ that we would get for a purely isotropic distribution. If the underlying distribution of HVGs is flattened to this extent, then the probability of matching all criteria rises to 41.7 ± 1.2‰.

However, we also need to account for the 25° mismatch between the axis about which the LG is flattened and the HVG plane normal. Although uncertainties of 10° are possible in the latter, they should be much smaller in the former due to the larger number of galaxies (35 instead of 8). Thus, it is difficult to explain our results as a consequence of our entire sample being flattened – it is, but about a different axis to the HVGs. Of course, our sample of LG galaxies may be subject to observational biases that make discovery more likely in certain directions on the sky. But the HVGs are likely subject to the same biases, unless there are selection effects based on the RV.

To quantify whether this misalignment is natural, we add the requirement[16] that the random 'flattening axis' $\widehat{\boldsymbol{n}}$ in each MC trial should be misaligned by at least as much as the observed 25° to the plane normal best fitting the mock HVGs. Including this additional requirement decreases the proportion of analogues to just 1.08 ± 0.02‰. This is not very surprising – if a MC trial 'flattened' along $\widehat{\boldsymbol{n}}$ yields a thin plane of mock HVGs with rms thickness lower than the observed 101 kpc, it probably does so because the normal to the mock HVG plane aligns rather closely with $\widehat{\boldsymbol{n}}$.

Because we include the HVGs when finding the properties of our entire sample, any plane preferred by the remaining galaxies would get tilted towards the plane preferred by the HVGs. To account for this, we exclude the HVGs and fit a plane to the remaining galaxies. Surprisingly, this yields an even more pronounced flattening, with the aspect ratio decreasing to 0.24. However, this also increases the misalignment between the flattening direction and the HVG plane normal. To see how these competing effects alter our results, we repeat our analysis with a more flattened prior distribution but require $\widehat{\boldsymbol{n}}$ to misalign with the mock HVG plane normal by at least 35°. This makes analogues to the observed HVG system even harder to find, with their frequency falling to just 0.12 ± 0.01‰.

7 DISCUSSION

Based on the criteria in Table 5, the observed system of HVGs appears most consistent with an underlying isotropic distribution if we use the aspect ratio (equation 35) instead of rms thickness (equation 31) and also remove Antlia from our sample while leaving HIZSS 3 in. Even with this highly favourable situation, only 5.35 ± 0.02‰ of the MC trials can be considered analogous to the observed situation (Table 7). Thus, it seems clear that an isotropic distribution of HVGs is very unlikely to mimic several aspects of the observed situation. This is a little unusual given that our full sample of galaxies does not prefer positions close to the plane defined by the HVGs (Fig. 11).

However, the existence of HVGs and their anisotropic distribution are less puzzling in the context of our proposed scenario where the high GRVs arose through three-body gravitational interactions of LG dwarfs with the MW and M31. Our simulation of this process (Section 3) implies that there should be a bimodal distribution of RVs at distances of ∼1.5–3 Mpc from the LG. Observationally, this is hinted at in Fig. 10 of this work as well as in fig. 9 of both Banik & Zhao (2016, 2017).

The process is likely to be more efficient for a LG dwarf flung out parallel to the motion of the major LG galaxy it interacted with. Thus, the fastest HVGs – which would now be furthest away from the LG – should preferentially lie close to the MW–M31 orbital plane, which of course also includes the MW–M31 barycentre (Fig. 9).

Even without detailed MC calculations, it is evident that most of the HVGs we identify (Table 3) lie rather close to a plane (Fig. 11). However, this is not true of our sample in general. There are dozens of galaxies (the vast majority) whose kinematics are consistent with expectations based on ΛCDM. These galaxies neither avoid nor prefer the plane defined by the HVGs, suggesting that their unusual kinematics is related to their anisotropic spatial arrangement.

Looking at the results of our statistical calculations (Table 7) in a less model-dependent way, one might focus on the probability of obtaining a plane thinner than observed (i.e. using just the first or second criterion in Table 5 and ignoring the others). In this case, our results are ∼4× larger. The highest result of 22.2 ± 0.1‰ arises when removing Antlia from our sample and using the aspect ratio instead of physical thickness.

The statistical significance of our results arises almost exclusively due to the galaxies in the NGC 3109 association, whose filamentary nature has been discussed previously (Bellazzini et al. 2013). Looking at fig. 1 of that work, it is clear that its conclusions are strengthened by a more recent distance measurement to Leo P (1.62 ± 0.15 Mpc, McQuinn et al. 2015). The rather high GRVs of

[16] To those listed in Table 5.

these galaxies were noted by Teyssier, Johnston & Kuhlen (2012) and Shaya & Tully (2013), the latter work finding no choice but to assume a past gravitational slingshot interaction with the MW ∼7 Gyr ago at a closest approach distance of 25 kpc (see their section 4.2.4). Although we consider this very likely, the role of dynamical friction in such close encounters was neglected, possibly missing an opportunity to discriminate between ΛCDM and modified gravity alternatives (Pawlowski & McGaugh 2014). Section 4.2 of that work reviews various explanations for the origin of both the spatial anisotropy and the kinematics of this association. However, the work does not consider the whole LG and lacks detailed dynamical modelling.

It is important to test if a scenario similar to what we propose for the LG might also have occurred elsewhere in the Universe. However, it is difficult to perform a timing argument analysis outside the LG because distance uncertainties are larger, making it harder to accurately pin down deviations from the Hubble flow (equation 25). None the less, the galaxy group containing NGC 1407 seems to have a much wider spread of RVs than is suggested by the variation in line of sight distances (Tully 2015). In particular, the galaxies NGC 1400 and NGC 1407 have a RV difference of nearly 1200 km s^{-1} even though it is quite likely that they are within 10 Mpc of each other, possibly much less (Tully et al. 2013). Several other galaxies in the group also show a rather wide spread in RV despite contamination issues not being so severe (Trentham, Tully & Mahdavi 2006). Although there is no clear evidence for a recent NGC 1400–NGC 1407 interaction (Spolaor et al. 2008), the galaxies are rather gas-poor ellipticals that could not be expected to long retain (or ever have) obvious features of a close interaction e.g. a recent starburst or tidal tails. Recent radio outbursts in the general vicinity are suggestive of the cluster gas sloshing around due to past motion of massive objects (Giacintucci et al. 2012). Deeper observations targeting more accurate distance measurements could help clarify the kinematics of this system.

7.1 Understanding our results in MOND

In our scenario, the HVGs should preferentially lie near a particular plane with normal along $\widehat{\boldsymbol{h}}$, the direction of the MW–M31 orbital angular momentum (Fig. 9). However, given the fairly small number of HVGs in our sample, it is possible for the observationally determined best-fitting plane to differ from any true underlying plane that is statistically preferred by the HVGs. This is because our MOND-based simulation of the LG (Section 3) indicates that not all the HVGs should orbit within exactly the same plane (Fig. 7).

Moreover, there must be observational errors in our determination of the HVG plane normal. Assuming distance errors of ∼0.1 Mpc for a structure ∼1 Mpc wide, this translates into an uncertainty of 6°, though the actual uncertainty is smaller as we have more than one HVG. There are also uncertainties in determining exactly which galaxies should be considered as HVGs. For example, removing Antlia from our sample tilts the HVG plane by 2.°5 (Table 4). This suggests observational errors of up to ∼10° in the HVG plane orientation.

Thus, we do not expect a perfect alignment between our estimated direction for $\widehat{\boldsymbol{h}}$ (Section 2.2) and that which best fits our sample of HVGs (Table 4). None the less, the 19° angle between these may be a problem for our scenario. In particular, the fact that the MW–M31 line is 16° out of the plane best fitting the HVGs could suggest that we are missing something. This angle is reduced to 15° if we remove Antlia from our sample of HVGs (Section 6.4), but is unlikely to be reduced to 0 given expected observational uncertainties.

However, there may well be deficiencies in our model which account for this discrepancy. One shortcoming is our treatment of the MW and M31 as point masses which have remained almost constant with time. We do not expect this to invalidate the main conclusion of our MOND model, namely that the particles flung out at high speeds by the MW/M31 tend to lie close to the MW–M31 orbital plane (Section 3). This is because encounters in MOND are similarly strong regardless of the impact parameter due to a cancellation between the encounter duration and the forces acting during the encounter. As a result, the vast majority of our simulated HVG analogues never passed within 5 disc scalelengths of either the MW or M31 (Fig. 3) − the commoner more distant encounters are strong enough.

These high-velocity particles gain most of their velocity in a cosmologically brief time period. Thus, our results should still hold if the MW and M31 masses can vary with time in a more realistic way. At present, it is unclear what a realistic MOND accretion history might look like. However, there is good reason to believe that structure formation would be more efficient than in ΛCDM (Llinares et al. 2008). This would mean that galaxies drain their surroundings efficiently at rather early times, to some extent justifying our 'island universe' approximation. Moreover, for a timing argument analysis, it is not important if gas 100 kpc from the MW gets accreted into its disc and triggers a burst of star formation. This would have significant consequences for the MW, but from the perspective of e.g. M31 at a distance of ∼800 kpc (McConnachie 2012), the gravitational attraction towards the MW would remain almost unchanged.

A potentially important effect beyond our point mass model is precession of the MW–M31 orbital pole $\widehat{\boldsymbol{h}}$ arising from the extended nature of their mass distributions and their close interaction. Due to the larger mass and disc scalelength of M31, the more important consideration is how the orbit of the MW may have precessed due to the non-spherical nature of the M31 gravitational field. The general direction of this precession can be estimated by applying equation (8) to the MW–M31 direction $\widehat{\boldsymbol{r}}_{\rm M31}$ at that time, which we find by rotating the present MW–M31 line about our best guess for $\widehat{\boldsymbol{h}}$ (Section 2.2). This suggests that $\widehat{\boldsymbol{r}}_{\rm M31}$ lay along Galactic coordinates (270°, 66°) at closest approach. Obtaining the M31 disc spin vector from Table 1, we find that $\widehat{\boldsymbol{h}}$ likely shifted towards the direction (334°, −11°) i.e. almost directly towards the orange line in Fig. 15, the locus of all directions where $\widehat{\boldsymbol{h}}$ could currently be given the direction towards M31. With a closest approach distance only a few times larger than the scalelength of the M31 disc, $\widehat{\boldsymbol{h}}$ could conceivably have precessed by ∼10°, enough to explain why $\widehat{\boldsymbol{r}}_{\rm M31}$ does not lie entirely within the HVG plane. More detailed models are required to account for such effects.

Another possibility is the effect of massive objects outside the LG. In Banik & Zhao (2016), we found that Centaurus A has a discernible impact on LG dynamics in a ΛCDM context. The stronger long-range gravitational force in MOND may enhance such effects, especially as the non-linearity of the theory means that even a constant external gravitational field on the LG can influence its internal dynamics. In such circumstances, the force exerted on a test particle by a massive object might not be directed towards it (e.g. Banik & Zhao 2015). Combined with tides, this can lead to tilting of a plane of LG galaxies defined by only a small number of them.

Such effects can be enhanced by the filamentary configuration of the HVGs within the plane they seem to define (Fig. 16). Taking an extreme example, suppose a Cartesian *xyz* coordinate system is used and that all the HVGs are located on the *xy*-plane along the lines L_1 and L_2 which are parallel to the *x*-axis. A perturber in the *yz*-plane at $z \neq 0$ would cause galaxies in L_1 to rise out of the *xy*-plane by

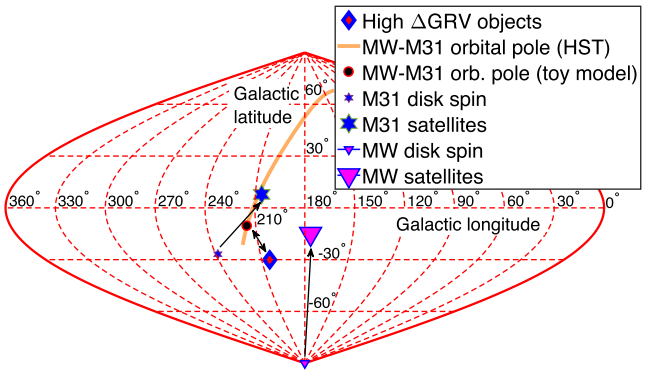


Figure 15. Directions of the vectors important to this work, in Galactic coordinates. Assuming a past close MW–M31 flyby, we expect tidal torque from M31 to explain the misalignment between the orientation of the MW disc (small triangle) and its plane of satellites (large triangle). The effect of such torques is illustrated using an upward arrow. Tidal torque from the MW explains a similar misalignment for M31 (hexagrams used instead of triangles). Our MOND-based toy model is able to reproduce these orientations fairly well if the MW–M31 orbital pole lies in the direction of the black dot with red rim (Section 2.2). This is reasonably consistent with the normal to the plane defined by the HVGs (diamond), though we do not yet know its sense of rotation. The proper motion of M31 has recently been measured (van der Marel et al. 2012b), suggesting a particular MW–M31 orbital pole (1σ allowed region shown as orange line). This must be orthogonal to the present direction towards M31. Unfortunately, at 2σ, any direction consistent with this requirement is allowed.

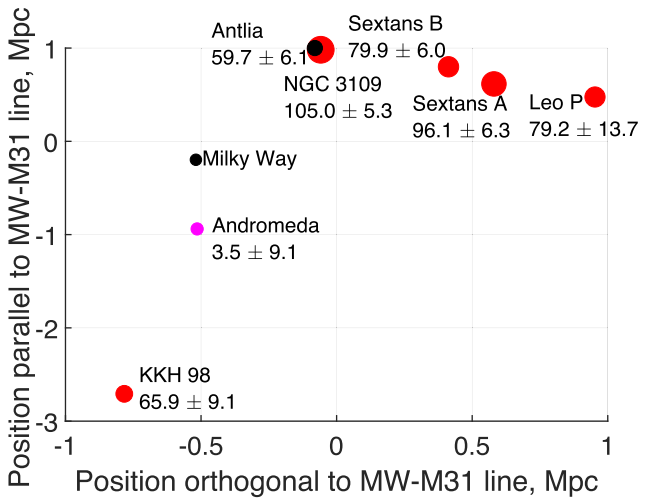


Figure 16. ΔGRVs of indicated galaxies are shown against their position within the best-fitting plane through them (orientation given in Table 4). Only HVGs close to this plane are shown, with the marker size proportional to ΔGRV (except for the MW and M31). The number below the name of each galaxy also shows its ΔGRV in km s^{-1}. For clarity, position uncertainties are not shown.

a different amount to galaxies in L_2, such that the best-fitting plane through all the HVGs would no longer be the xy-plane. The analysis of such effects is beyond the scope of this investigation, but may be feasible based on purely geometric arguments.

7.2 Understanding our results in ΛCDM

It is possible that there is a ΛCDM-based explanation for the results discussed in this contribution. However, the anomalously high GRVs of the HVGs are unlikely to be explained by the tidal effects of large-scale structure (Banik & Zhao 2017, section 5.1) or by massive galaxy groups just outside the LG – these are already directly included in our dynamical model (see its table 3). Although that analysis could have missed dynamical solutions involving close interactions outside the LG, this is only likely to be a viable solution for distant galaxies like NGC 4163 which have an extremely low GRV compared to expectations (Fig. 10). It is conceivable that this galaxy was flung towards the LG by a close interaction outside it which is not properly included in our simulation. However, there is insufficient time for a galaxy to be flung towards the LG and now be moving away from the LG again (Banik & Zhao 2017, section 5.2). Moreover, our improvements to that analysis (Section 4.1) make it less likely that we miss such encounters.

Our ΛCDM-based timing argument did not allow for the masses of galaxies changing with time, similar to previous works (e.g. Peñarrubia et al. 2014). Although galaxies are expected to grow with time (e.g. Wechsler et al. 2002), this should not much affect our results for several reasons. The main one is that the timing argument is mostly sensitive to late times because an impulse at earlier times would change both the present position and velocity of a particle. This causes an effect similar to but steeper than Hubble drag (Banik & Zhao 2016, Fig. 4). A reduction in masses at earlier times would also need to be compensated by a higher present mass in order to get a similar time-averaged gravitational field in the LG and thus to match observed GRVs. Moreover, for e.g. the MW–M31 gravitational attraction to be substantially affected by accretion on to the MW, the accreted material would need to come from a very large distance. Even if it came from just beyond the MW virial radius ∼200 kpc away (Dehnen, McLaughlin & Sachania 2006), this is still only a small fraction of the ∼800 kpc distance to M31. Thus, at the large scales investigated in this contribution, it should be acceptable to treat LG galaxies as having a fixed mass.

Given that our model handles tides raised by objects outside the LG and the paucity of galaxies with an unusually low ΔGRV (Fig. 10), the most likely explanation for the HVGs is that they were flung out by massive fast-moving object(s) *inside* the LG several Gyr ago. Our investigation does not elucidate the nature of these object(s), tempting as it may be to identify them with the MW and M31. This is infeasible in ΛCDM as the theory implies no past MW–M31 flyby – in any close interaction, dynamical friction between their dark matter haloes would cause a merger (Privon et al. 2013). Without a past interaction, the MW and M31 would always have been slow-moving relative to the LG barycentre, limiting their ability to account for the HVGs.[17]

Moreover, the growth of the MW and M31 masses with time implies that their scattering power must have been smaller at earlier times (Wechsler et al. 2002). Indeed, MW and M31 analogues in ΛCDM simulations do not appear to fling out dark matter haloes beyond ∼3 virial radii and even these 'associated haloes' have rather small outwards velocities (Sales et al. 2007, figs 3 and 6).

Dynamical friction is less efficient for less massive objects like M33 and the Large Magellanic Cloud, but both of these are already included in our model at velocities consistent with their observed

[17] This is not true at very early times, when the Hubble expansion was very fast. However, impulses at those times hardly affect present motions due to Hubble drag. We demonstrated this explicitly by showing that our results hardly changed when we started our simulations earlier (Banik & Zhao 2016, section 4.6).

GRVs and proper motions (Banik & Zhao 2017, Table 3). A galaxy with even less mass would have an even smaller circular velocity (Evrard et al. 2008), making it unlikely that it could ever have scattered the HVGs on to their presently observed orbits.

Given the significance of dynamical friction for heavier galaxies, it is possible that another galaxy X merged with e.g. M31, building up a high velocity relative to it by falling deep into its potential well. Any LG dwarf passing close to the space–time location of the X-M31 interaction could then be flung out at high speed, taking up some of the kinetic energy of X if X was sufficiently massive. However, too high a mass could disrupt the M31 disc and limit the X-M31 relative velocity due to more significant dynamical friction. On the other hand, too low a mass is also not feasible due to the need to scatter the HVGs we identified, some of which have rather high masses. For example, NGC 3109 is rotating at $\sim\frac{1}{3}$ the rate of the MW so must have a few percent of its mass (Jobin & Carignan 1990). Sextans A and B also have substantial masses of $\sim 10^9 \mathrm{M}_\odot$ (Bellazzini et al. 2014). These considerations may leave a range of plausible masses for X, though we argued previously that the effects of Hubble drag imply it would only have enough scattering power to explain our results if its mass was comparable to that of the MW or M31 (Banik & Zhao 2017, section 5.2).

In such a scenario, the past gravitational field in the LG would have been different to that assumed in our ΛCDM timing argument calculation. However, this is mostly sensitive to forces acting at late times, thus limiting the effect of an error in the gravitational field at early times (Banik & Zhao 2017, fig. 4). If we really have got the past gravitational field in the LG wrong at a recent enough time to affect our results, then one might expect some galaxies to have GRVs much below the predictions of our (erroneous) model. Our analysis indicates that this very rarely happens (Fig. 11). Moreover, it is difficult to find analogues to the HVGs in cosmological simulations of ΛCDM that include realistic merger histories for analogues of the MW and M31 (Sales et al. 2007, fig. 2).

A related scenario is that the HVGs were formed by tidal disruption of X, making them tidal dwarf galaxies flung out during the merger of a gas-rich galaxy X with a major LG galaxy. The chaotic motions and complicated gas hydrodynamics during this interaction may provide a way around the fact that the slow motion of the MW and the even slower motion of M31 do not readily provide a mechanism to create HVGs. One consequence of this scenario is that the HVGs should have rather low internal velocity dispersions because tidal dwarf galaxies should be almost free of dark matter (Wetzstein et al. 2007). However, the internal dynamics of Tucana (Fraternali et al. 2009) and NGC 3109 (Jobin & Carignan 1990) indicate that they require dark matter for dynamical stability in a ΛCDM context. This is also the case for Sextans A and B (Bellazzini et al. 2014).

The dark matter in the HVGs could be understood if X had its own retinue of satellite galaxies. The X-M31 interaction would disrupt this satellite system, perhaps creating a few HVGs. In one such scenario, X is identified with NGC 205 and its satellites have now formed a structure analogous to a tidal stream around M31, helping to explain its anisotropic distribution of satellites (Angus et al. 2016).

Assuming that some satellites of NGC 205 could escape from M31, an obvious consequence of this scenario should be that the HVGs lie in the same plane as that preferred by satellites of M31. However, there is a 41° mismatch between the orientations of these planes (Fig. 15), far larger than likely observational uncertainties. The plane of M31 satellites may have precessed from its original orientation (Fernando et al. 2017), but such a large amount of precession would almost certainly inflate the thickness of the structure as not all M31 satellites would be equally affected.

Although satellites of NGC 205 could plausibly end up at the fairly low velocities typical of M31 satellites, it is questionable whether they could become HVGs. Using a test particle cloud around NGC 205 to represent its satellites, it is clear that only a minuscule fraction (if any) of these particles end up further than 200 kpc from M31 at the present time (Angus et al. 2016, fig. 12). This is because a fairly low infall velocity is required to ensure that a substantial number of NGC 205 satellites become bound to M31.

As the mass of NGC 205 which worked best in these models was only 1 per cent that of M31 itself (Angus et al. 2016, section 2.2.2), any satellites of NGC 205 would likely end up in a bound orbit around M31, just like NGC 205 itself. Applying the $M \propto v_\mathrm{f}^3$ scaling typical of ΛCDM haloes (Evrard et al. 2008) and assuming $v_\mathrm{f} = 225\,\mathrm{km\,s^{-1}}$ for M31 (Carignan et al. 2006), we get $\sim$45 $\mathrm{km\,s^{-1}}$ for the typical velocity of a NGC 205 satellite relative to its host. This is much smaller than the difference between the circular and escape velocities of M31, which must be at least $(\sqrt{2}-1)v_\mathrm{f} \approx 90\,\mathrm{km\,s^{-1}}$. The actual figure could be far higher due to the extended nature of the M31 mass distribution. For example, the MW escape velocity 50 kpc from it is $\sim$380 $\mathrm{km\,s^{-1}}$ (Williams et al. 2017) despite v_f only being $\sim$200 $\mathrm{km\,s^{-1}}$ for the MW (Kafle et al. 2012). Thus, it is easy to see why this scenario does not lead to a satellite of NGC 205 ending up as far from M31 as the HVGs.

We should bear in mind that this proposal was not designed to explain our HVGs. For this purpose, a variant could be considered where X is no longer identified with NGC 205 but instead fell towards M31 with a much higher relative velocity. This might well lead to a filamentary structure receding from the LG at high speed, with X perhaps identified as the NGC 3109 association. However, this does not explain the high infall velocity of X, which seems difficult to reconcile with ΛCDM (Banik & Zhao 2017, Section 5.2).

Moreover, a fairly massive dark matter halo would be needed to hold the NGC 3109 association together (Bellazzini et al. 2013). It would have to pass quite close to the MW/M31 in order to get tidally disrupted, as would be required for the galaxies in this association with end up in their observed filamentary configuration. It is unclear whether dynamical friction would render such a scenario infeasible, though that appears quite likely if the association has a width of 600 kpc and a mass of $3.2 \times 10^{11}\,\mathrm{M}_\odot$ (as suggested by Bellazzini et al. 2013). To account for the rather high GRVs of the galaxies in this association, it needs to have passed within 25 kpc of the MW (Shaya & Tully 2013, section 4.2.4) without subsequently merging or being significantly decelerated.

In addition to this issue, several other difficulties with the scenario were discussed in section 4.2.1 of Pawlowski et al. (2014). Moreover, the latest results indicate that the galaxies in this association are very unlikely to be gravitationally bound and probably lie within their own individual dark matter haloes (Kourkchi & Tully 2017, section 9). The only exception is Antlia, which could be bound to NGC 3109. However, excluding Antlia from our HVG sample does not greatly alter our conclusions (Section 6.4). If the remaining HVGs are not gravitationally bound to each other, then it is reasonable to consider these dwarf galaxies as independent. In this case, our results show that they have a statistically significant tendency to lie close to a plane (Table 7). Thus, it remains difficult to simultaneously explain the high RVs of the HVGs and their anisotropic spatial distribution within the ΛCDM paradigm.

8 CONCLUSIONS

We recently identified several LG galaxies with much higher RVs than can easily be understood in the context of ΛCDM (Banik & Zhao 2016, 2017). These HVGs are not bound to the MW or M31 but instead lie $\gtrsim 1$ Mpc away from them. This is $\gtrsim 5\times$ the MW virial radius (Dehnen et al. 2006), well beyond the furthest distance to which cosmological simulations indicate that the MW or M31 could slingshot out dark matter haloes containing LG dwarf galaxies (Sales et al. 2007, figs 3 and 6). The issue therefore arises on a much larger scale than that addressed in previous works regarding the anisotropic distribution of the MW and M31 satellite systems (e.g. Pawlowski et al. 2014). In a ΛCDM context, baryonic physics does not seem to have a major effect on the expected anisotropy of these systems (Pawlowski et al. 2015, fig. 3). Thus, it seems unlikely that it would affect our results on a still larger scale.

In this contribution, we assess the feasibility of a scenario involving a past close flyby of the MW and M31, whose once fast relative motion could have flung out these HVGs via gravitational slingshot encounters. To help constrain the orientation of the MW–M31 orbital plane, we develop a toy model of their flyby forming the recently discovered planes of satellites around the MW (Pawlowski & Kroupa 2013) and M31 (Ibata et al. 2013). A past close MW–M31 encounter seems able to form correctly oriented satellite planes only for a particular MW–M31 orbital pole (Section 2.2).

Using this information, we constructed a MOND-based dynamical model of the LG to investigate the effect of a past MW–M31 flyby on it (Section 3). We tracked the evolution of a spherical cloud of several hundred thousand LG test particles initially on the Hubble flow, leaving a gap around the MW and M31 (equation 27). Although our results are somewhat dependent on details of how MOND works in a cosmological context, it is clear that a small fraction of these particles end up with a large RV away from the LG after closely interacting with the MW or M31 around the time of their encounter (Fig. 5). Such slingshot interactions are expected to be most efficient for particles flung out almost parallel to the motion of the perturbing body. This probably explains why simulated particles flung out to the greatest distances from the LG preferentially lie close to the MW–M31 orbital plane (Fig. 9).

To see if such a pattern is evident in the real LG, we made several improvements to our ΛCDM modelling of the LG (Section 4.1) and developed a method for selecting HVGs (Section 4.2) so that we could quantify their spatial anisotropy (Section 5). The HVGs we identified are indeed mostly located close to a plane (Fig. 11) oriented similarly to our expectation for the MW–M31 orbital plane based on the flyby scenario (Fig. 15). The galaxies in our sample which are not HVGs (the vast majority) do not preferentially lie close to the HVG plane. Importantly, the MW and M31 lie near this plane (barycentre $\sim$70 kpc off it), suggesting that it has dynamical significance.

These characteristics of the observed HVG plane are natural outcomes of our MOND simulation of the LG (Section 3). To quantify how likely it is that a random distribution of HVGs shares these properties, we conducted MC trials where the directions to the HVGs were randomized and their distances selected from a Gaussian distribution corresponding to observations (Section 6).[18]

Based on the criteria in Table 5, there is a <0.01 probability of obtaining a situation analogous to that observed, mainly because it is very unusual for all but one of the HVGs (listed in Table 3) to so nearly lie within a common plane. Our result holds even if we vary certain assumptions, with the nominal choices leading to a still smaller value of only $1.5 \pm 0.1‰$ (Table 7). As the observed HVG plane has an aspect ratio of 0.08 (Table 4), our results are not consistent with a mildly flattened prior distribution for the HVGs (Section 6.6). More extreme flattening could lead to consistency and may well be realistic as our entire sample of LG galaxies does indeed exhibit a moderate flattening. However, the preferred flattening axis is misaligned by 25°[19] to the HVG plane normal. Thus, the flattened distribution of the HVGs is very likely related to their anomalous RVs and not merely a consequence of the entire LG being moderately anisotropic about a different axis.

As well as more detailed modelling, our ideas regarding the LG may be tested using a more accurate distance to HIZSS 3 to see if it really should be treated as a HVG. Better observations of Antlia could help clarify whether it is a satellite of NGC 3109 or an independent galaxy. In the long run, it is important to try and discover more HVGs to see if they mostly remain close to a plane.[20] If so, it will be interesting to know whether its orientation is altered much compared to our determination. Considering the effect of Antlia on this, adjustments of up to $\sim 5^\circ$ could well be in store (Table 4).

Our work suggests a particular great circle on the sky in which observers would be more likely to discover HVGs. However, it is important to search other directions too as only then is it possible to determine just how anisotropic the distribution of HVGs is. Fortunately, the plane we identified is inclined by a large angle to the MW disc and to the Ecliptic, which should help minimize observational biases. Moreover, there are unlikely to be any strong selection effects based on the RV.

Both 2D and 3D ΛCDM-based dynamical models of the LG face difficulties in explaining the observed kinematics of its non-satellite galaxies (Banik & Zhao 2016, 2017). In particular, the existence of seven galaxies with anomalously high RVs but only one with anomalously low RV (Fig. 10) suggests that the central region of the LG has been much more efficient at scattering dwarf galaxies than these models allow. This is problematic given that even LG galaxies as massive as the MW and M31 appear incapable of scattering dwarfs as far out as the HVGs, not only in our simulations but also when considered in full cosmological simulations of ΛCDM (Sales et al. 2007). This may hint that the MW and M31 were once moving much faster than at present, pointing towards a past close MW–M31 flyby. If such an event occurred, it would provide a natural explanation for several aspects of the spatial distribution of the HVGs, especially their tendency to lie close to a plane (Figs 9 and 11). A past flyby interaction of such massive galaxies only makes sense in the context of certain modified gravity theories where galaxies lack massive dark matter haloes and their associated dynamical friction in close encounters, which would otherwise cause a merger.

ACKNOWLEDGEMENTS

The authors are grateful to Marcel Pawlowski for suggesting the idea behind this contribution. IB is supported by Science and Technology Facilities Council studentship 1506672. This work was partly

[18] We converted negative mock distances to 0.

[19] 35° when excluding HVGs.

[20] Observers do not search for HVGs per se but instead for LG dwarf galaxies which may turn out to be HVGs when their kinematics are analysed using a ΛCDM-based dynamical model.

supported by a Scottish Universities Physics Alliance travel award hosted by Nima Arkani-Hamed at the Institute for Advanced Studies. The algorithms were set up using MATLAB.

REFERENCES

Angus G. W., Coppin P., Gentile G., Diaferio A., 2016, MNRAS, 462, 3221
Aragon-Calvo M. A., Silk J., Szalay A. S., 2011, MNRAS, 415, L16
Banik I., 2014, preprint (arXiv:1406.4538v2)
Banik I., Zhao H., 2015, preprint (arXiv:1509.08457)
Banik I., Zhao H., 2016, MNRAS, 459, 2237
Banik I., Zhao H., 2018, MNRAS, 473, 419
Banik I., Zhao H., 2017, MNRAS, 467, 2180
Barnes D. G., de Blok W. J. G., 2001, AJ, 122, 825
Bekenstein J., Milgrom M., 1984, ApJ, 286, 7
Bell E. F., de Jong R. S., 2001, ApJ, 550, 212
Bellazzini M., Oosterloo T., Fraternali F., Beccari G., 2013, A&A, 559, L11
Bellazzini M. et al., 2014, A&A, 566, A44
Belokurov V. et al., 2014, MNRAS, 437, 116
Bovy J., Rix H.-W., 2013, ApJ, 779, 115
Brimioulle F., Seitz S., Lerchster M., Bender R., Snigula J., 2013, MNRAS, 432, 1046
Butsky I. et al., 2016, MNRAS, 462, 663
Caldwell N. et al., 2017, ApJ, 839, 20
Candlish G. N., Smith R., Fellhauer M., 2015, MNRAS, 446, 1060
Carignan C., Chemin L., Huchtmeier W. K., Lockman F. J., 2006, ApJ, 641, L109
Cautun M., Bose S., Frenk C. S., Guo Q., Han J., Hellwing W. A., Sawala T., Wang W., 2015, MNRAS, 452, 3838
Chemin L., Carignan C., Foster T., 2009, ApJ, 705, 1395
Chiboucas K., Jacobs B. A., Tully R. B., Karachentsev I. D., 2013, AJ, 146, 126
Corbelli E., Salucci P., 2007, MNRAS, 374, 1051
Courteau S., Widrow L. M., McDonald M., Guhathakurta P., Gilbert K. M., Zhu Y., Beaton R. L., Majewski S. R., 2011, ApJ, 739, 20
Cyburt R. H., Fields B. D., Olive K. A., Yeh T.-H., 2016, Rev. Mod. Phys., 88, 015004
Dalcanton J. J. et al., 2009, ApJS, 183, 67
Dehnen W., McLaughlin D. E., Sachania J., 2006, MNRAS, 369, 1688
den Heijer M. et al., 2015, A&A, 581, A98
Desmond H., 2017, MNRAS, 464, 4160
Einasto J., Lynden-Bell D., 1982, MNRAS, 199, 67
Evans I. N. et al., 2010, ApJS, 189, 37
Evrard A. E. et al., 2008, ApJ, 672, 122
Famaey B., Binney J., 2005, MNRAS, 363, 603
Famaey B., McGaugh S. S., 2012, Living Rev. Relativ., 15, 10
Fernando N., Arias V., Guglielmo M., Lewis G. F., Ibata R. A., Power C., 2017, MNRAS, 465, 641
Fletcher R., Powell M. J. D., 1963, Comput. J., 6, 163
Fraternali F., Tolstoy E., Irwin M. J., Cole A. A., 2009, A&A, 499, 121
Giacintucci S. et al., 2012, ApJ, 755, 172
Gilmore G., Reid N., 1983, MNRAS, 202, 1025
González R. E., Padilla N. D., 2010, MNRAS, 407, 1449
Hayden M. R. et al., 2015, ApJ, 808, 132
Ibata R. A. et al., 2013, Nature, 493, 62
Ibata R. A., Ibata N. G., Lewis G. F., Martin N. F., Conn A., Elahi P., Arias V., Fernando N., 2014, ApJ, 784, L6
Iocco F., Pato M., Bertone G., 2015, Phys. Rev. D, 92, 084046
Jacobs B. A., Rizzi L., Tully R. B., Shaya E. J., Makarov D. I., Makarova L., 2009, AJ, 138, 332
Jayaraman A., Gilmore G., Wyse R. F. G., Norris J. E., Belokurov V., 2013, MNRAS, 431, 930
Jobin M., Carignan C., 1990, AJ, 100, 648
Jurić M. et al., 2008, ApJ, 673, 864
Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2012, ApJ, 761, 98
Kahn F. D., Woltjer L., 1959, ApJ, 130, 705
Karachentsev I. D., Makarov D. I., Kaisina E. I., 2013, AJ, 145, 101
Keller B. W., Wadsley J. W., 2017, ApJ, 835, L17
Keller B. W., Wadsley J., Couchman H. M. P., 2016, MNRAS, 463, 1431
Kormendy J., Drory N., Bender R., Cornell M. E., 2010, ApJ, 723, 54
Kourkchi E., Tully R. B., 2017, ApJ, 843, 16
Kroupa P., 2015, Can. J. Phys., 93, 169
Kroupa P., Theis C., Boily C. M., 2005, A&A, 431, 517
Lehner N., Howk J. C., Wakker B. P., 2015, ApJ, 804, 79
Lelli F., McGaugh S. S., Schombert J. M., 2016, AJ, 152, 157
Lelli F., McGaugh S. S., Schombert J. M., Pawlowski M. S., 2017, ApJ, 836, 152
Llinares C., Knebe A., Zhao H., 2008, MNRAS, 391, 1778
Longmore A. J., Hawarden T. G., Webster B. L., Goss W. M., Mebold U., 1978, MNRAS, 183, 97P
López-Corredoira M., Kroupa P., 2016, ApJ, 817, 75
Lüghausen F., Famaey B., Kroupa P., 2015, Can. J. Phys., 93, 232
Lynden-Bell D., 1976, MNRAS, 174, 695
Lynden-Bell D., 1982, Observatory, 102, 202
Ma J., 2001, Chin. Phys. Lett., 18, 1420
Ma C. et al., 1998, AJ, 116, 516
Massey P., Henning P. A., Kraan-Korteweg R. C., 2003, AJ, 126, 2362
McConnachie A. W., 2012, AJ, 144, 4
McConnachie A. W. et al., 2009, Nature, 461, 66
McGaugh S. S., 2011, Phys. Rev. Lett., 106, 121303
McGaugh S. S., 2016a, ApJ, 816, 42
McGaugh S. S., 2016b, ApJ, 832, L8
McGaugh S., Milgrom M., 2013, ApJ, 775, 139
McGaugh S. S., Schombert J. M., 2015, ApJ, 802, 18
McGaugh S., Lelli F., Schombert J., 2016, Phys. Rev. Lett., 117, 201101
McMillan P. J., 2017, MNRAS, 465, 76
McQuinn K. B. W. et al., 2015, ApJ, 812, 158
Metz M., Kroupa P., Jerjen H., 2007, MNRAS, 374, 1125
Metz M., Kroupa P., Jerjen H., 2009, MNRAS, 394, 2223
Milgrom M., 1983, ApJ, 270, 365
Milgrom M., 1999, Phys. Lett. A, 253, 273
Milgrom M., 2010, MNRAS, 403, 886
Milgrom M., 2013, Phys. Rev. Lett., 111, 041105
Milgrom M., 2016, preprint (arXiv:1610.07538)
Mirabel I. F., Dottori H., Lutz D., 1992, A&A, 256, L19
Müller O., Jerjen H., Pawlowski M. S., Binggeli B., 2016, A&A, 595, A119
Müller O., Scalera R., Binggeli B., Jerjen H., 2017, A&A, 602, A119
Nicastro F., Mathur S., Elvis M., 2008, Science, 319, 55
Nicastro F., Senatore F., Krongold Y., Mathur S., Elvis M., 2016, ApJ, 828, L12
Noh Y., Lee J., 2007, ApJ, 666, 627
Norris M. A. et al., 2016, ApJ, 832, 198
Pawlowski M. S., Kroupa P., 2013, MNRAS, 435, 2116
Pawlowski M. S., McGaugh S. S., 2014, MNRAS, 440, 908
Pawlowski M. S., Kroupa P., de Boer K. S., 2011, A&A, 532, A118
Pawlowski M. S., Pflamm-Altenburg J., Kroupa P., 2012, MNRAS, 423, 1109
Pawlowski M. S. et al., 2014, MNRAS, 442, 2362
Pawlowski M. S., Famaey B., Merritt D., Kroupa P., 2015, ApJ, 815, 19
Pazy E., 2013, Phys. Rev. D, 87, 084063
Peebles P. J. E., 2017, preprint (arXiv:1705.10683)
Peebles P. J. E., Nusser A., 2010, Nature, 465, 565
Peebles P. J. E., Tully R. B., Shaya E. J., 2011, preprint (arXiv:1105.5596)
Peñarrubia J., Ma Y.-Z., Walker M. G., McConnachie A., 2014, MNRAS, 443, 2204
Penny S. J., Pimbblet K. A., Conselice C. J., Brown M. J. I., Grützbauch R., Floyd D. J. E., 2012, ApJ, 758, L32
Piatek S., Pryor C., Bristow P., Olszewski E. W., Harris H. C., Mateo M., Minniti D., Tinney C. G., 2006, AJ, 131, 1445
Pimbblet K. A., Couch W. J., 2012, MNRAS, 419, 1153
Planck Collaboration XXVI 2014, A&A, 571, A27
Planck Collaboration XIII 2016, A&A, 594, A13
Privon G. C., Barnes J. E., Evans A. S., Hibbard J. E., Yun M. S., Mazzarella J. M., Armus L., Surace J., 2013, ApJ, 771, 120

Quillen A. C., Garnett D. R., 2001, in Funes J. G., Corsini E. M., eds, ASP Conf. Ser. Vol. 230, Galaxy Disks and Disk Galaxies. Astron. Soc. Pac., San Francisco, p. 87
Raychaudhury S., Lynden-Bell D., 1989, MNRAS, 240, 195
Riess A. G. et al., 1998, AJ, 116, 1009
Salem M., Besla G., Bryan G., Putman M., van der Marel R. P., Tonnesen S., 2015, ApJ, 815, 77
Sales L. V., Navarro J. F., Abadi M. G., Steinmetz M., 2007, MNRAS, 379, 1475
Sandage A., 1986, ApJ, 307, 1
Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103
Shaya E. J., Tully R. B., 2013, MNRAS, 436, 2096
Silva D. R., Massey P., DeGioia-Eastwood K., Henning P. A., 2005, ApJ, 623, 148
Smolin L., 2017, preprint (arXiv:1704.00780)
Snaith O. N., Haywood M., Di Matteo P., Lehnert M. D., Combes F., Katz D., Gómez A., 2014, ApJ, 781, L31
Spolaor M., Forbes D. A., Hau G. K. T., Proctor R. N., Brough S., 2008, MNRAS, 385, 667
Swaters R. A., Sancisi R., van Albada T. S., van der Hulst J. M., 2009, A & A, 493, 871
Teyssier R., 2002, A&A, 385, 337
Teyssier M., Johnston K. V., Kuhlen M., 2012, MNRAS, 426, 1808
Thies I., Kroupa P., Famaey B., 2016, preprint (arXiv:1606.04942)
Tikhonov A. V., Klypin A., 2009, MNRAS, 395, 1915
Tiret O., Combes F., 2008, in Funes J. G., Corsini E. M., eds, ASP Conf. Ser. Vol. 396, Formation and Evolution of Galaxy Disks. Astron. Soc. Pac., San Francisco, p. 259
Torrealba G., Koposov S. E., Belokurov V., Irwin M., 2016, MNRAS, 459, 2370
Trentham N., Tully R. B., Mahdavi A., 2006, MNRAS, 369, 1375
Tully R. B., 2015, AJ, 149, 54
Tully R. B., Fisher J. R., 1977, A&A, 54, 661
Tully R. B. et al., 2013, AJ, 146, 86
van den Bergh S., 1999, ApJ, 517, L97
van der Marel R. P., Fardal M., Besla G., Beaton R. L., Sohn S. T., Anderson J., Brown T., Guhathakurta P., 2012a, ApJ, 753, 8
van der Marel R. P., Besla G., Cox T. J., Sohn S. T., Anderson J., 2012b, ApJ, 753, 9
Verlinde E. P., 2016, preprint (arXiv:1611.02269)
Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691
Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52
Wetzstein M., Naab T., Burkert A., 2007, MNRAS, 375, 805
Williams A. A., Belokurov V., Casey A. R., Evans N. W., 2017, MNRAS, 468, 2359
Yin J., Hou J. L., Prantzos N., Boissier S., Chang R. X., Shen S. Y., Zhang B., 2009, A & A, 505, 497
Zhao H., Famaey B., Lüghausen F., Kroupa P., 2013, A&A, 557, L3

This paper has been typeset from a TEX/LATEX file prepared by the author.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **473,** 419–430 (2018)
Advance Access publication 2017 September 11

doi:10.1093/mnras/stx2350

The escape velocity curve of the Milky Way in modified Newtonian dynamics

Indranil Banik★ and Hongsheng Zhao

Scottish Universities Physics Alliance, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9SS, UK

Accepted 2017 September 6. Received 2017 September 6; in original form 2017 August 12

ABSTRACT

We determine the escape velocity from the Milky Way (MW) at a range of Galactocentric radii in the context of modified Newtonian dynamics (MOND). Due to its non-linear nature, escape is possible if the MW is considered embedded in a constant external gravitational field (EF) from distant objects. We model this situation using a fully self-consistent method based on a direct solution of the governing equations out to several thousand disc scalelengths. We try out a range of EF strengths and mass models for the MW in an attempt to match the escape velocity measurements of Williams et al. (2017). A reasonable match is found if the EF on the MW is $\sim 0.03 a_0$, towards the higher end of the range considered. Our models include a hot gas corona surrounding the MW, but our results suggest that this should have a very low mass of $\sim 2 \times 10^{10}\,\mathrm{M}_\odot$ to avoid pushing the escape velocity too high. Our analysis favours a slightly lower baryonic disc mass than the $\sim 7 \times 10^{10}\,\mathrm{M}_\odot$ required to explain its rotation curve in MOND. However, given the uncertainties, MOND is consistent with both the locally measured amplitude of the MW rotation curve and its escape velocity over Galactocentric distances of 8−50 kpc.

Key words: gravitation – stars: kinematics and dynamics – Galaxy: disc – Galaxy: fundamental parameters – Galaxy: kinematics and dynamics – large-scale structure of Universe.

1 INTRODUCTION

The standard Λ cold dark matter (ΛCDM) cosmological paradigm (Ostriker & Steinhardt 1995) has a great deal of flexibility in fitting the rotation curves (RCs) of individual galaxies due to the unknown relation between their baryonic content and their often dominant dark matter (DM) distribution required by this model (e.g. Rubin et al. 1980). This makes it difficult to extract unique RC predictions from ΛCDM, as illustrated by de Blok & McGaugh (1998). Their fig. 6 shows that ΛCDM can fit the RC of NGC 2403 rather well based on the photometry data of UGC 128, a different galaxy with a much lower surface brightness and indeed a rather different RC. In this model, it should therefore be difficult to predict the RCs of individual galaxies based solely on their baryonic distribution.

However, observations indicate the opposite (Famaey & McGaugh 2012, and references therein). In a ΛCDM context, spiral galaxy RCs exhibit a tight correlation between their shape, dark matter halo scale radius and m0ass such that $\lesssim 10^{-5}$ of the available phase space volume is actually filled (Salucci et al. 2007). Those authors noted that 'theories of the formation of spirals do not trivially imply the existence of such a surface that underlies the occurrence of a strong dark-luminous coupling'. Although it has long been known that the Newtonian gravity $\boldsymbol{g}_N$ of the baryons is insufficient to *explain* the RC of many galaxies, the RC can nonetheless be *predicted* from $\boldsymbol{g}_N$ alone by scaling it in a universally valid way. This radial acceleration relation (RAR) has recently been clarified with space-based *spitzer* observations in the near-infrared (McGaugh, Lelli & Schombert 2016), taking advantage of the lower dispersion in mass-to-light ratios at these wavelengths (Bell & de Jong 2001; Norris et al. 2016). Similar analyses are sometimes possible in elliptical galaxies, especially those that contain a thin rotation-supported gas disc (den Heijer et al. 2015). The RAR works extremely well in both classes of galaxy over ∼5 orders of magnitude in luminosity and a similar range of surface brightness (Lelli et al. 2017). Deviations from the RAR fall within the expected observational uncertainties and appear to be uncorrelated with any of the numerous parameters that could plausibly be relevant, e.g. surface brightness and gas fraction (see their fig. 4).

We consider the RAR to be more natural if galaxies are not surrounded by DM haloes but their purported dynamical effect instead arises instead from an acceleration-dependent modification to Newtonian gravity, a hypothesis called Modified Newtonian Dynamics (MOND, Milgrom 1983). This assumes that the gravitational field strength g at distance r from an isolated point mass M transitions from the usual inverse square law at short range to

$$g = \frac{\sqrt{GMa_0}}{r} \quad \text{for} \quad r \gg \sqrt{\frac{GM}{a_0}}. \tag{1}$$

★ E-mail: ib45@st-andrews.ac.uk

Here, a_0 is a fundamental acceleration scale of nature which must have an empirical value close to 1.2×10^{-10} m s^{-2} to match galaxy rotation curves (McGaugh 2011). In a remarkable coincidence called the cosmic coincidence of MOND, a_0 is comparable to the value of g at which a classical gravitational field has an energy density equal to the dark energy density $u_\Lambda = \rho_\Lambda c^2$ implied by the accelerating expansion of the Universe (Riess et al. 1998). Thus,

$$\frac{g^2}{8\pi G} < u_\Lambda \quad \Leftrightarrow \quad g \lesssim 2\pi a_0. \tag{2}$$

This suggests that MOND may be caused by quantum gravity effects (e.g. Milgrom 1999; Pazy 2013; Verlinde 2016; Smolin 2017). Regardless of its underlying microphysical explanation, it works well at explaining the dynamics of isolated galaxies. In the Local Group (LG), it requires that the Milky Way (MW) and M31 have undergone a past close flyby (Zhao et al. 2013) due to their strong gravitational attraction in MOND and the almost radial nature of the MW-M31 orbit (van der Marel et al. 2012). Such a flyby is not possible in ΛCDM because dynamical friction between their DM haloes would inevitably cause a merger (Privon et al. 2013). However, it would provide a natural explanation for several LG galaxies with very high radial velocities (RVs), much higher than that can easily be accounted for in a 3D dynamical model of the LG in ΛCDM (Banik & Zhao 2017a). These RVs remain anomalously high despite several improvements to the model and the procedure used to find its best-fitting parameters (Banik & Zhao 2017b, section 4.1). However, their fig. 5 shows that the high-velocity galaxies (HVGs) have RVs broadly consistent with the speeds at which a once fast-moving MW or M31 could have gravitationally slingshot out LG dwarfs in three-body gravitational interactions governed by MOND.

One consequence of the MOND model should be that the HVGs lie rather close to the MW-M31 orbital plane. This is because it should be easier to scatter a dwarf galaxy to a very high RV if it is scattered parallel to the motion of the perturber. We recently used a test particle model to demonstrate this and also showed that the observed spatial distribution of the HVGs is indeed rather anisotropic (Banik & Zhao 2017b).

A past MW-M31 interaction might also have formed the thick disc of the MW (Gilmore & Reid 1983), a structure which formed fairly rapidly from its thin disc 9 ± 1 Gyr ago (Quillen & Garnett 2001). More recent investigations confirm a fairly rapid formation timescale (Hayden et al. 2015) and an associated burst of star formation (Snaith et al. 2014, fig. 2). The disc heating which likely formed the Galactic thick disc appears to have been stronger in the outer parts of the MW, characteristic of a tidal effect (Banik 2014). This may be why the thick disc of the MW has a larger scalelength than its thin disc (Jurić et al. 2008; Jayaraman et al. 2013).

In this contribution, we test a more subtle consequence of MOND called the external field effect (EFE) which arises because the theory is acceleration-dependent (Milgrom 1986, section 2g). To understand it, consider a dwarf galaxy with low internal accelerations ($\ll a_0$) freely falling in the strong acceleration ($\gg a_0$) of a distant massive galaxy such that there are no tidal effects. The overall acceleration at any point in the dwarf is rather high due to the dominant external field (EF) of the massive galaxy. Thus, the dwarf would obey Newtonian dynamics and forces in its vicinity would follow the usual inverse square law rather than equation (1). However, without the massive galaxy, the internal dynamics of the dwarf would be very non-Newtonian.

Using the principle of continuity, the RC of a galaxy must be slightly affected even if the EF on it is much weaker than its internal gravity. Applying this idea, Haghi et al. (2016) analysed whether the RCs of a sample of 18 disc galaxies could be fit better in MOND once the EFE is considered. Their work relied on a plausible analytic estimate of how the EFE would weaken the internal gravity of these galaxies. In most of the cases considered, non-zero values of the EF were preferred due to the RCs declining faster than expected in the outer regions if one neglects the EFE. Moreover, the preferred EF strengths were roughly consistent with the expected gravity from other known galaxies in the vicinity of the 18 they considered (see their fig. 7).

Perhaps the clearest demonstration of the EFE is in the velocity dispersion of the MW satellite Crater 2, which was predicted to be only $2.1^{+0.6}_{-0.3}$ km s^{-1} in MOND (McGaugh 2016b, section 4.1). The rather low value is partly due to the EFE of the nearby MW, without which the prediction would have been $\sim$4 km s^{-1}. This is in tension with the observed value of 2.7 ± 0.3 km s^{-1} (Caldwell et al. 2017). Thus, the internal dynamics of Crater 2 are not consistent with a naive application of the RAR but are consistent with a more rigorous treatment of MOND and its inevitable EFE.

At large distances from an object, the EF is likely to be the dominant source of gravity. We previously derived the far-field MOND forces generated by a point mass embedded in a constant and dominant EF of magnitude $g_{\rm ext}$ (Banik & Zhao 2015). g eventually transitions to an inverse square law with a super-Newtonian normalization if $g_{\rm ext} \ll a_0$, as will be the case in this work. Thus, a point mass should create a potential well of finite depth even in MOND. Moreover, we showed that the escape velocity would differ by $\lesssim$3 *per cent* between the most common versions of the theory, namely the aquadratic Lagrangian formulation (AQUAL, Bekenstein & Milgrom 1984) and the quasi-linear formulation (QUMOND, Milgrom 2010). QUMOND is much simpler to handle numerically because it only requires a solution to the normal Poisson equation, albeit twice as often as for Newtonian gravity.[1] This is much simpler than the non-linear grid relaxation method required in AQUAL (Brada & Milgrom 1999), so we will focus on QUMOND in this contribution. This is also the basis for the publicly available Phantom of RAMSES algorithm (Lüghausen, Famaey & Kroupa 2015) which 'MONDifies' the gravity law in the RAMSES *N*-body hydrodynamics solver (Teyssier 2002). A similar algorithm has recently been developed that can solve AQUAL as well, though this is not yet public (Candlish, Smith & Fellhauer 2015).

Realizingthat MOND with the EFE predicts potential wells of finite depth, Famaey, Bruneton & Zhao (2007) used an analytic method to estimate the escape velocity $v_{\rm esc}$ from the MW in the vicinity of the Sun. Similar results were later obtained by Wu et al. (2008) using a numerical solution to AQUAL. Their estimated $v_{\rm esc}$ agrees reasonably well with later measurements based on high-velocity MW stars (Piffl et al. 2014). Recently, a similar technique was used to measure $v_{\rm esc}$ over a wide range of Galactocentric radii (8−50 kpc, Williams et al. 2017). This work applied the method of Leonard & Tremaine (1990) to a variety of tracers detected in the ninth data release of the Sloan Digital Sky Survey (SDSS, Ahn et al. 2012). We wish to calculate the expected $v_{\rm esc}$ in MOND at these positions.

To better explore the range of plausible MW mass models, we include a hot gas halo surrounding its stellar and gas discs. This is suggested by X-ray spectroscopic observations at a range of Galactic

[1] The increased computational cost is offset against the fact that QUMOND should work without the addition of dark matter particles, at least on galactic scales.

latitudes (Nicastro et al. 2016) and by the truncation of the Large Magellanic Cloud (LMC) gas disc, most likely a consequence of ram pressure stripping (Salem et al. 2015). A similar halo has recently been detected around M31 based on quasar sightline observations (Lehner, Howk & Wakker 2015).

In Section 2, we explain the method by which we accurately determine $v_{\rm esc}$ in our MOND models of the MW. The results thus obtained are shown in Section 3 and their accuracy is discussed in Section 4, where we also consider other issues such as the plausibility of our best-fitting model parameters based on independent considerations. Our conclusions are given in Section 5.

2 METHOD

QUMOND uses the Newtonian gravitational field $\boldsymbol{g}_{\rm N}$ due to a matter distribution as the first of two stages in calculating the gravitational field $\boldsymbol{g}$:

$$\overbrace{\nabla\cdot\boldsymbol{g}}^{\propto\rho_{\rm PDM}+\rho_b} \equiv \nabla\cdot\left[\nu\left(\frac{|\boldsymbol{g}_{\rm N}|}{a_0}\right)\boldsymbol{g}_{\rm N}\right], \quad \text{where} \tag{3}$$

$$\nu\left(x\right) = \frac{1}{2} + \sqrt{\frac{1}{4}+\frac{1}{x}}. \tag{4}$$

Here, we use the 'simple' interpolating function $\nu(x)$ to go between the Newtonian and low-acceleration regimes (Famaey & Binney 2005). $\nabla\cdot\boldsymbol{g}$ is the source term for the gravitational field, so it can be thought of as an 'effective' density ρ composed of the actual density $\rho_{\rm b}$ and an extra term which we define to be the phantom dark matter density $\rho_{\rm PDM}$. This is the distribution of dark matter that would be necessary in Newtonian gravity to generate the same gravitational field as in QUMOND. The Newtonian gravitational field $\boldsymbol{g}_{\rm N} \equiv -\nabla\Phi_{\rm N}$ satisfies the usual Poisson equation

$$\nabla^2\Phi_{\rm N} = 4\pi G\rho_{\rm b}. \tag{5}$$

After solving this with the usual isolated boundary conditions ($g\rightarrow 0$ as $r\rightarrow\infty$), we add the contribution from the Newtonian EF $\boldsymbol{g}_{\rm N,ext}$, which is what the EF on the MW would have been in Newtonian gravity. We assume the spherically symmetric MOND relation between this and the actual EF $\boldsymbol{g}_{\rm ext}$ on the MW:

$$\Phi \rightarrow \Phi - \boldsymbol{r}\cdot\boldsymbol{g}_{\rm N,ext}, \quad \text{where} \tag{6}$$

$$\overbrace{\nu\left(\frac{g_{\rm N,ext}}{a_0}\right)}^{\nu_{\rm ext}}\boldsymbol{g}_{\rm N,ext} = \boldsymbol{g}_{\rm ext}. \tag{7}$$

2.1 The Newtonian potential

The MW is assumed to consist of a hot gas corona surrounding two aligned and concentric infinitely thin exponential discs representing its gas and stellar components. Taking advantage of the fact that we can superpose potentials in Newtonian gravity, we simply add the potential of the corona to that of the other components. The corona is treated as a Plummer model (Plummer 1911) with mass $M_{\rm cor}$ and core radius $r_{\rm cor}$, yielding a corona potential at distance r from the MW of

$$\Phi_{\rm cor} = -\frac{GM_{\rm cor}}{\sqrt{r^2+{r_{\rm cor}}^2}}. \tag{8}$$

For the disc components, we only need to solve for a single exponential disc. We take this to have unit scalelength and GM so that it can be scaled up to the required values later. We discretize equation (5) and solve it in spherical polar co-ordinates (polar angle θ) using the successive overrelaxation method described in Appendix A.

Once we obtain Φ in this way, we scale it up to the correct mass and length scale for the MW stellar disc and superpose another scaled version of this solution to represent its gas disc. We then add the corona potential (equation 8) and the external field (equation 6).

2.2 The QUMOND gravitational field

Using the discretisation scheme described in Appendix B, we determine the QUMOND source density $\nabla\cdot\boldsymbol{g} \equiv \nabla\cdot\left(\nu\boldsymbol{g}_{\rm N}\right)$. We integrate this directly in order to obtain $\boldsymbol{g}$ itself, which we only need at a small fraction of the grid points:

$$\boldsymbol{g}\left(\boldsymbol{r}\right) = \int \nabla\cdot\boldsymbol{g}(\boldsymbol{r}')\frac{(\boldsymbol{r}-\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3}\,{\rm d}^3\boldsymbol{r}'. \tag{9}$$

For axisymmetric problems, the forcing $\nabla\cdot(\nu\boldsymbol{g}_{\rm N})$ can be considered as a large number of uniform density rings, making it simple to determine $\boldsymbol{g}$ on the symmetry axis via direct summation. For off-axis points, we avoid an excessive computational cost by making use of a 'ring library' which stores the gravity exerted by a thin ring with $GM_{\rm ring} = r_{\rm ring} = 1$, where $M_{\rm ring}$ and $r_{\rm ring}$ refer to the ring mass and radius, respectively. To find the gravity at any point due to a ring, we scale the relative co-ordinates to the appropriate position within our ring library and interpolate to obtain the required result before scaling by $\frac{GM_{\rm ring}}{{r_{\rm ring}}^2}$ at the end. We neglect rings passing very close to the point on which we calculate $\boldsymbol{g}$ as these rather large contributions should almost completely cancel, but it is difficult to handle such cancellation accurately on a computer. We consider the effect of slightly different 'excluded regions' and use cubic extrapolation to estimate $\boldsymbol{g}$ if there had been no excluded region at all.

We can only consider contributions to $\boldsymbol{g}$ from a finite volume. Our aim is to consider a sufficiently large volume that contributions from more distant regions can be handled analytically. In particular, we cover a large enough region that the EF should start dominating beyond it. This transition occurs at a distance of

$$r_{\rm ext} = \frac{\sqrt{GMa_0}}{g_{\rm ext}}. \tag{10}$$

In the models we consider (Table 1), $r_{\rm ext}$ can be as distant as 1220 kpc. Thus, we use our method to obtain $\nabla\cdot\boldsymbol{g}$ out to a distance $r_{\rm out} = 13\,341$ kpc. The phantom dark matter in the EF-dominated region beyond this is not spherically symmetric and thus contributes to the gravity at smaller distances. We account for this assuming EF dominance in the regions beyond $r_{\rm out}$, allowing us to analytically determine the gravity resulting from phantom dark matter there as that follows the density distribution derived in Banik & Zhao (2015, section 3). This leads to an adjustment of

$$\boldsymbol{g} \rightarrow \boldsymbol{g} + (3\cos^2\theta - 1)\frac{GM\nu_{\rm ext}\widehat{\boldsymbol{r}}}{15{r_{\rm out}}^3}, \quad \text{where} \quad \widehat{\boldsymbol{r}} \equiv \frac{\boldsymbol{r}}{|\boldsymbol{r}|}. \tag{11}$$

Assuming that we are considering a point where the EF is dominant, the gravity due to the MW has a magnitude of $\sim\frac{GM\nu_{\rm ext}}{r^2}$ such that the correction to it expressed in fractional terms is only $\sim\frac{1}{15}\left(\frac{r}{r_{\rm out}}\right)^3$. Thus, the accuracy of our results should not depend much on this correction, which should in any case be fairly accurate as it estimates contributions from regions with $r > 11r_{\rm ext}$. There, the Newtonian gravitational field due to the MW should be $>120\times$ weaker than $g_{\rm N,ext}$, allowing it to be considered perturbatively.

Table 1. Our adopted parameters for the MW mass distribution, with a $_0$ subscript indicating a nominal value. We always use the same value of $\frac{M_g}{M_*}$ and the same disc scalelengths, but vary the other parameters using a grid search. The first part of our table contains the fixed parameters $R_\odot$ (McMillan 2017), $M_{*,0}$ (McGaugh 2016a), r_* (Bovy & Rix 2013), r_g (McMillan 2017) and $M_{g,0}$, which is based on applying the method described in McGaugh (2008, section 3.3) to the observations of Olling & Merrifield (2001, table D1). The Galactic hot gas corona is modelled using equation (14) (Plummer 1911).

Variable	Meaning	Value
$R_\odot$	Galactocentric distance of Sun	8.2 kpc
r_*	Stellar disc scalelength	2.15 kpc
$M_{*,0}$	Nominal stellar disc mass	$5.51 \times 10^{10}\,M_\odot$
r_g	Gas disc scalelength	7 kpc
$M_{g,0}$	Nominal gas disc mass	$1.18 \times 10^{10}\,M_\odot$
$\frac{M_*}{M_{*,0}}$	disc mass scaling factor	0.8–1.4
r_{cor}	Plummer radius of corona	(20–60) kpc
M_{cor}	Corona mass	$(2–8) \times 10^{10}\,M_\odot$
g_{ext}	External field on MW	$(0.01–0.03)\,a_0$

To minimize edge effects of the sort just mentioned, we only use our method to obtain $\boldsymbol{g}$ out to a distance of 3892 kpc, where equation (11) requires us to correct $\boldsymbol{g}$ by $\sim\frac{1}{600}$. The EF is not really dominant at this point, so some method is required to estimate $\boldsymbol{g}$ at larger distances. Only then can we determine $\boldsymbol{g}$ far enough out into the EF-dominated region for us to 'hand over' to the analytic results found by Banik & Zhao (2015). For this purpose, we construct another library holding the gravitational field due to a point mass embedded in a constant EF assuming the deep-MOND limit (DML).[2] It is not totally accurate to assume the DML because the EF is still a few per cent of a_0 (Famaey et al. 2007). Thus, we adjust forces in this region by the ratio of ν_{ext} to its value ν_{DML} assuming the DML:

$$\frac{\nu}{\nu_{DML}} = 1 + \frac{g_{ext}}{a_0}, \quad \text{where} \tag{12}$$

$$\nu_{DML} \equiv \sqrt{\frac{a_0}{g_{N,ext}}}. \tag{13}$$

To see how appropriate a point mass model is for the MW, we need to consider its most extended component – its hot gas corona. The fraction of its total mass enclosed within any radius r is given by

$$\frac{M\,(<r)}{M} = \frac{r^3}{\left(r^2 + {r_{cor}}^2\right)^{\frac{3}{2}}}. \tag{14}$$

In the most extended case we consider, $r_{cor} = 60$ kpc so that only $\sim 10^{-4}$ of the corona lies at radii beyond 3.9 Mpc. Having thus verified the accuracy of our point mass and DML assumptions and corrected their deficiencies as far as possible, we use them to obtain the gravitational field out to $r = 66.5 r_{ext}$. Here, we find that the forces are within a few per cent of the analytic results of Banik & Zhao (2015) which we therefore use to obtain the depth of the potential well at this point:

$$\Phi = -\frac{GM\nu_{ext}}{r}\left(1 + \frac{K_0}{2}\sin^2\theta\right), \quad \text{where} \tag{15}$$

[2] To avoid recalculating it for different EF strengths.

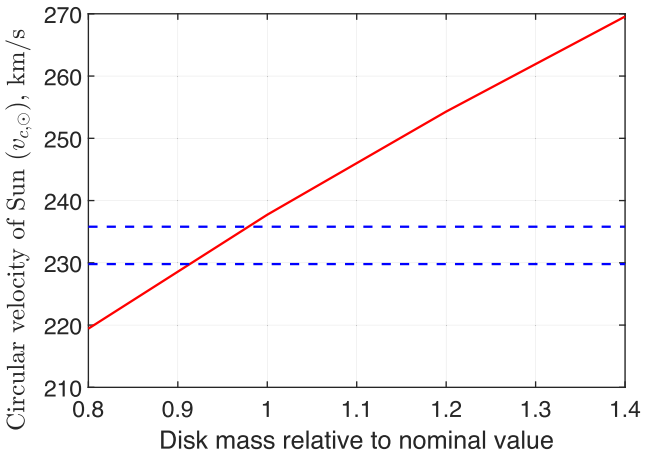


Figure 1. The effect on $v_{c,\odot}$ of scaling our nominal stellar and gas disc masses (Table 1) by the amount shown. The dashed blue lines show the 1σ allowed range for $v_{c,\odot}$ (McMillan 2017).

$$K_0 \equiv \frac{\partial Ln\ \nu_{ext}}{\partial Ln\ g_{N,ext}} = -\frac{1}{2} \quad \text{if} \quad g_{N,ext} \ll a_0. \tag{16}$$

To summarize, our potential calculations are based on considering the gravitational field in three regions. The innermost one covers out to 3892 kpc and considers the MW mass distribution in detail as well as the EF it is embedded in, without assuming anything about how strong the resulting gravity is. In the next region covering out to $66.5 r_{ext}$, the MW is treated as a point mass in an EF without assuming the gravity from either is dominant. Although this aspect of the problem is assumed to be in the DML, a small correction is applied to account for g_{ext} being a few per cent of a_0. The outermost region uses the analytic potential arising from a point mass embedded in a dominant EF (equation 15).

3 RESULTS

Our model for the MW mass distribution is designed to be consistent with its observed rotation curve in a MOND context (McGaugh 2016a, table 1 model Q4ZB). This particular model does not require a bulge to get the gravitational field strength correct in the Solar neighbourhood because of its large distance from the Galactic Centre and the rather short length scale associated with any bulge component. Indeed, McMillan (2017) suggest that 80 per cent of its mass lies within just 2.2 kpc (see their section 2.1) whereas the Sun is nearly four times more distant (see their table 2). At even larger distances, the dynamical effect of the bulge should be even smaller and easily accounted for with minor adjustments to the disc normalizations and scalelengths. Moreover, their work also contains a central hole in the MW gas disc whereas our model does not, partly compensating for our lack of a bulge component. Thus, our results should not be greatly affected by this assumption.

We begin by considering how our four model parameters (Table 1) influence $v_{c,\odot}$. This depends mainly on the disc surface density such that only the nominal model is able to reproduce this correctly if we assume that $v_{c,\odot} \approx 235\,\mathrm{km\,s^{-1}}$ (McMillan 2017). However, we consider the effect of scaling the surface density by factors of 0.8–1.4 (Fig. 1).

Adjusting M_{cor} can affect $v_{c,\odot}$ by $\lesssim 5\,\mathrm{km\,s^{-1}}$ while adjusting its scalelength r_{cor} has a smaller effect of $\sim 1\,\mathrm{km\,s^{-1}}$. At the Solar circle, the MW is effectively isolated – raising g_{ext} from the lowest to

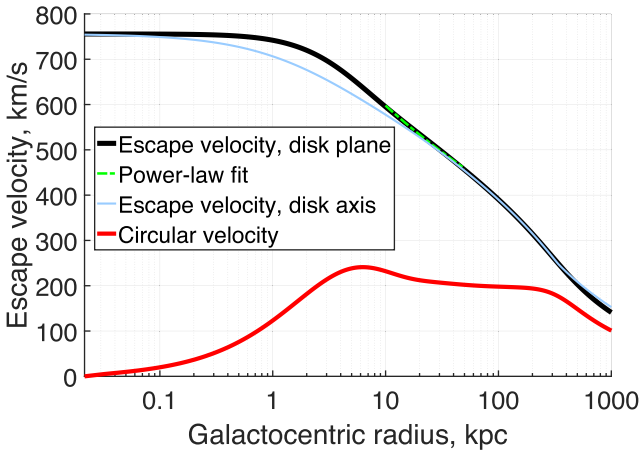


Figure 2. How the circular velocity of the MW (lower red curve) and its escape velocity (upper black curve) depend on position within its disc plane. The latter can be parametrized rather well as a power law (equation 17) over the radial range 10–50 kpc (dashed green curve). At the same distance from the MW, its escape velocity is lower along its disc axis (thin blue curve) for points close to the MW due to the effect of its disc. However, this pattern is reversed at long range because we assume the EF on the MW is aligned with its disc axis, deepening the potential in this direction (Banik & Zhao 2015, equation 37). The model shown here uses the nominal disc masses in Table 1 and $g_{\rm ext} = 0.03a_0$, with the corona being as small and low mass as possible.

the highest values considered only affects $v_{\rm c,\odot}$ by $\sim$4 m s^{-1}. These factors must have a more significant influence on forces further from the MW, but the scarcity of tracers makes it difficult to directly measure $\boldsymbol{g}$ there. Fortunately, forces at large r affect the escape velocity $v_{\rm esc} = \sqrt{-2\Phi}$ near the Sun, which must therefore be rather more sensitive to these parameters. This gives us the opportunity to constrain the MW gravitational field at large distances based on observations relatively close to it.

Using our nominal values for the MW stellar and gas disc masses (Table 1), we determined its circular and escape velocity curves within its disc plane (Fig. 2). For comparison with observations, we fit a power-law model to $v_{\rm esc}$ over the radial range 10–50 kpc. This assumes that

$$v_{\rm esc}(r) \propto r^{-\alpha}. \tag{17}$$

Power-law fits become linear when considering the logarithms of both variables. Thus, if we let $y \equiv \mathrm{Ln}\, v_{\rm esc}$ and $x \equiv \mathrm{Ln}\, r$ be lists of size N, then we have that

$$\tilde{x} \equiv x - \frac{1}{N}\sum_{i=1}^{N} x_i \quad (\tilde{y} \text{ defined analogously}) \tag{18}$$

$$\alpha = -\frac{\sum_{i=1}^{N} \tilde{x}_i \tilde{y}_i}{\sum_{i=1}^{N} \tilde{x}_i \tilde{x}_i}. \tag{19}$$

We can fit $v_{\rm esc}(r)$ rather well using a power law over the range $r =$ 10–50 kpc (Fig. 2). For a range of models, we compare the resulting slope α and normalization at the Solar circle to observations (Figs 3 and 4). It is unclear exactly which Galactic polar angles θ the observations of Williams et al. (2017) correspond to, but most likely a range of angles is used in order to get enough of the relatively rare high-velocity stars that are necessary for an escape velocity determination. Thus, we show results for points along the disc symmetry axis ($\theta = 0$) in Fig. 3 and within the disc plane ($\theta = \frac{\pi}{2}$) in Fig. 4. The escape velocities are slightly larger in the latter case because the MW matter distribution gets closer to a point within its disc plane than an equally distant point along its disc axis. Within $\sim$100 kpc of the MW, this near-field effect is more important than the non-sphericity of the MW potential in the far-field EF-dominated region (equation 15), where the MW exerts very little gravity in any case. However, beyond $\sim$100 kpc, the latter effect dominates because the MW can be considered as a point mass. This leads to a deeper potential in the direction of $\boldsymbol{g}_{\rm ext}$, which we take to be aligned with the disc axis for simplicity (Fig. 2).

The results for $\theta = 0$ are hardly affected if we set $\boldsymbol{g}_{\rm ext} \rightarrow -\boldsymbol{g}_{\rm ext}$ because the forces at long range are symmetric with respect to $\theta \rightarrow \pi - \theta$ (equation 15). Naturally, they are also symmetric at short range as the EF is unimportant here and the MW model is symmetric. This minimizes any scope for $v_{\rm esc}$ differing between equally distant points along $\theta = 0$ and π, or equivalently at the same point but with $\boldsymbol{g}_{\rm ext} \rightarrow -\boldsymbol{g}_{\rm ext}$. Although our algorithm is unable to rigorously consider intermediate EF orientations, this should have only a small effect on our results (Section 4.3).

4 DISCUSSION

Our calculations for the MW's escape velocity are broadly consistent with the observations of Williams et al. (2017) over the Galactocentric distances they cover (8–50 kpc), especially when considering that the Solar neighbourhood escape velocity could be as high as 690 km s^{-1} at the 95 per cent confidence level (see their fig. 3). Moreover, it is probably easier to underestimate $v_{\rm esc}$ than to overestimate it. The velocity distribution of stars is expected to drop off close to $v_{\rm esc}$, but it is easy to imagine the theoretical distribution function not being filled all the way up to $v_{\rm esc}$ because it is impossible to have less than one star. This could well be why earlier studies underestimated v_{esc} locally. For instance, Meillon et al. (1998) found a 90 per cent confidence upper limit of just 550 km s^{-1} despite using *Hipparcos* data (Perryman 1989) supplemented by accurate radial velocities (Udry et al. 1997). Even if the positions and velocities of all MW stars were known perfectly, we could quite plausibly find none moving faster than $0.9v_{\rm esc}$ (Smith et al. 2007, fig. 1).

This issue should have been alleviated somewhat by the efforts of Williams et al. (2017) to constrain the precise form of the cut-off in the velocity distribution close to $v_{\rm esc}$ (see their section 5.1). Nonetheless, it must persist at some level and likely becomes more severe further from the MW due to the lower stellar number density. This can lead to a faster apparent decline in $v_{\rm esc}$ than is actually the case, perhaps explaining their rather high inferred value of α (equation 17) compared to our models.

Another issue might be that stars do not need to get infinitely far from the MW in order to escape. If they get to 400 kpc, then this is half-way to M31 in some directions (McConnachie 2012). At this point, our calculations suggest that $v_{\rm esc}$ is still $\sim$150 km s^{-1} (Fig. 2). Subtracting this from 600 km s^{-1} in quadrature suggests that the local escape velocity could be reduced by $\sim$20 km s^{-1} due to the presence of M31. This effect would be larger 50 kpc from the MW, where $v_{\rm esc}$ is only $\sim$450 km s^{-1}. Thus, $v_{\rm esc}(r)$ might well decline faster than in our models. This could raise α by $\sim$0.02, possibly more if M31 is heavier than the MW, as appears likely (Peñarrubia et al. 2014; Banik & Zhao 2016). Such effects would depend on the direction relative to the direction towards M31, necessitating a fully 3D model for the MW because M31 does not lie very close to the Galactic disc plane. In this case, it would be important to carefully consider the survey volume to better understand how M31 might influence $v_{\rm esc}$. This is beyond the scope of our analysis and probably not worthwhile given the present uncertainties on $v_{\rm esc}$, but

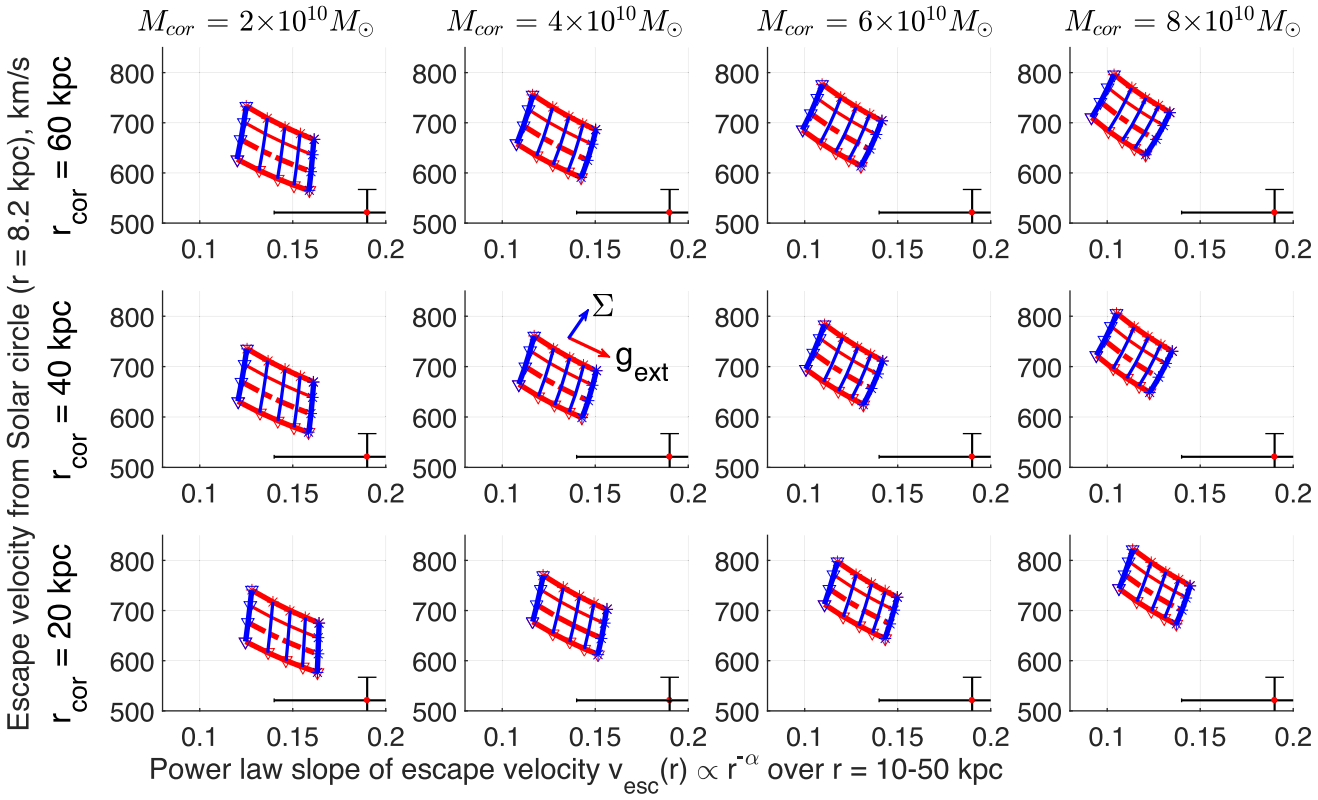


Figure 3. Escape velocity v_{esc} from within the MW disc plane as a function of the model parameters. The x-axis shows the value of α such that $v_{esc}(r) \propto r^{-\alpha}$ while the y-axis shows v_{esc} near the Sun. The measured values of these quantities are shown as a red dot with black error bars towards the bottom right (Williams et al. 2017). Each subplot has a fixed corona mass and scalelength, with red tracks showing the effect of varying g_{ext} with constant disc mass (vice versa for blue tracks). In each case, an inverted triangle is used to show the result when the parameter being varied has the lowest value we consider while a star is used for the largest value. We consider disc masses scaled from their nominal values by factors given in Table 1, where we also show the range in g_{ext} that we try (values of all parameters are spaced linearly). The dashed red lines show the results for the nominal stellar and gas disc masses, which are required to obtain the correct $v_{c,\odot}$ (Fig. 1). We assume $\boldsymbol{g_{ext}}$ is aligned with the disc symmetry axis.

may become important in future when more accurate measurements become available.

One indication that these effects do not greatly influence the analysis of Williams et al. (2017) comes from a relation between the circular and escape velocity curves that arises because both are determined by the same potential Φ:

$$\frac{\partial\left(\frac{1}{2}{v_{esc}}^2\right)}{\partial r} = -\frac{\partial\Phi}{\partial r} \tag{20}$$

$$= -\frac{{v_c}^2}{r} \tag{21}$$

$$\alpha \equiv -\frac{\partial Ln\, v_{esc}}{\partial Ln\, r} = \left(\frac{v_c}{v_{esc}}\right)^2. \tag{22}$$

If we assume that $v_{c,\odot} = 232.8\,\mathrm{km\,s^{-1}}$ (McMillan 2017), then we expect $\alpha = 0.200$ for a local escape velocity of $521\,\mathrm{km\,s^{-1}}$. This is entirely consistent with the observed value of 0.19 ± 0.05. The mild tension could indicate that α is indeed 0.2 rather than 0.19, though our analysis suggests that the solution lies instead with higher v_{esc}.

Our models treat the disc components of the MW as infinitely thin. This seems a reasonable approximation given the rather small scaleheights of the MW thin and thick discs (Snaith et al. 2014). Still, the disc components have a finite thickness, yielding a shallower potential well within the disc plane. This would bring our model more in line with observations if they are primarily sensitive to low Galactic latitudes (Fig. 3). If instead they are more sensitive towards the Galactic poles, then the effect of thickening the disc is likely smaller. This is fortunate as our model predictions are already more consistent with observations if they are along the disc axis (Fig. 4).

Our results are somewhat uncertain due to imperfect knowledge of $R_\odot$. Presently, this is constrained to within ∼0.1 kpc (Chatzopoulos et al. 2015; McMillan 2017). Increasing $R_\odot$ from our adopted 8.2 to 8.5 kpc reduces v_c and v_{esc} in the Solar neighbourhood by ∼1 and $3\,\mathrm{km\,s^{-1}}$, respectively. These effects are much smaller than the observational uncertainties, especially for v_{esc}.

The uncertain LSR speed $v_{c,\odot}$ also has some effect on our results. The analysis of McMillan (2017) found that $v_{c,\odot}$ could be as low as $220\,\mathrm{km\,s^{-1}}$ if the present constraints on $R_\odot$ are not considered. In this case, MOND requires a lower surface density Σ for the MW to match a slower rotation curve. Fig. 1 suggests that we might need to scale Σ by ∼0.8, thereby reducing v_{esc} near the Solar neighbourhood by $\sim 40\,\mathrm{km\,s^{-1}}$ while leaving α almost unaltered (Fig. 3).

However, such models rely on a very low value of $v_{c,\odot}$ which in turn implies a rather low $R_\odot$ given tight constraints on the ratio of these quantities from the proper motion of the supermassive black hole at the centre of the MW (Brunthaler et al. 2007). This is in significant tension with independent measurements of $R_\odot$. Once these are considered, it becomes clear that $v_{c,\odot}$ is constrained to be $232.8 \pm 3.0\,\mathrm{km\,s^{-1}}$, making a value of $220\,\mathrm{km\,s^{-1}}$ highly unlikely (McMillan 2017). It is thus difficult to improve our results

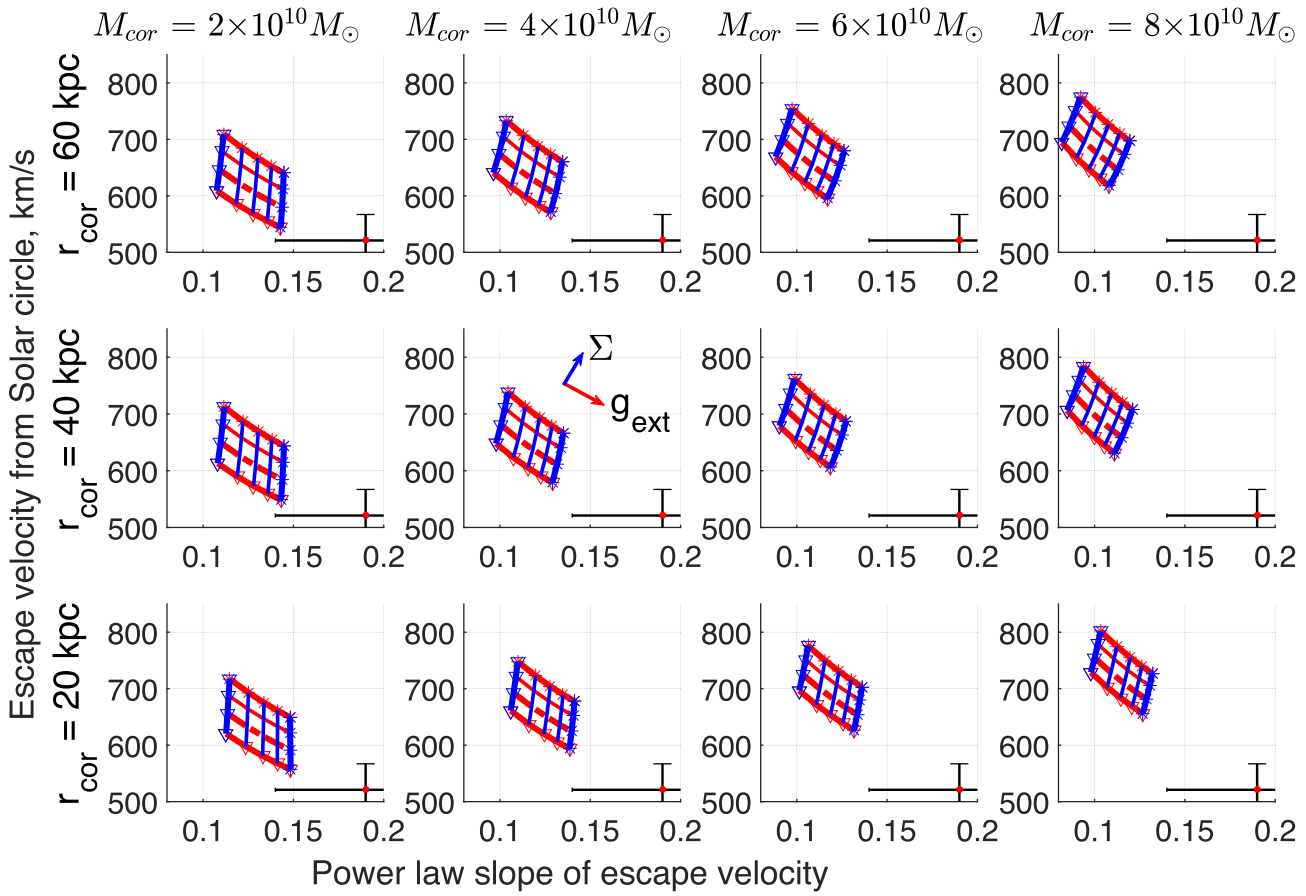


Figure 4. Similar to Fig. 3 but for v_{esc} along the disc symmetry axis in the direction of $\boldsymbol{g}_{ext}$. We obtained very similar results when going in the opposite direction (not shown), presumably because the radial gravity at both short and long range (isolated or EF-dominated) is the same at points with polar angles of θ and $\pi - \theta$.

significantly through tighter constraints on the position and velocity of the Sun with respect to the MW. Instead, we suggest that observers should focus on improving measurements of v_{esc}.

Our results shed new light on the issue of whether the LMC (LMC) is bound to the MW. At a Galactocentric distance of 50 kpc (Pietrzyński et al. 2013), we expect that $v_{esc} \approx 440\,\mathrm{km\,s^{-1}}$ for an MW disc having the mass required to explain its circular velocity curve in a MOND context (McGaugh 2016a, table 1 model Q4ZB). If we scale the disc mass down by 0.8 (lowest red tracks in Figs 3 and 4), then v_{esc} would fall by $\approx$30 km s^{-1}. Even then, 410 km s^{-1} greatly exceeds the Galactocentric velocity of the LMC as this is only $321 \pm 24\,\mathrm{km\,s^{-1}}$ (Kallivayalil et al. 2013, table 5). Thus, the LMC is almost certainly a bound satellite of the MW in a MOND context unless the EF on it significantly exceeds the maximum value of $0.03a_0$ considered here, a conclusion also reached by Wu et al. (2008). Even without considering dynamical models of the MW, the observations of Williams et al. (2017) alone indicate that $v_{esc} = 379^{+34}_{-28}\,\mathrm{km\,s^{-1}}$ at the distance of the LMC, strongly suggesting that it is bound to the MW.

Although our analysis is based on standard MOND, this tends to underestimate the forces binding galaxy clusters (e.g. Sanders 2003). One possible solution is Extended MOND (EMOND, Zhao & Famaey 2012), which posits that the acceleration scale a_0 increases with the potential depth $|\Phi|$ such that it has the standard value of $1.2 \times 10^{-10}\,\mathrm{m\,s^{-2}}$ for $|\Phi| \ll \Phi_0$ but is larger in regions with a deeper potential, such as galaxy clusters. So far, this appears to be a promising way to resolve the difficulties typically faced by MOND in such systems (Hodson & Zhao 2017). The required value of Φ_0 corresponds to a speed of $\sqrt{2\Phi_0} = 1800\,\mathrm{km\,s^{-1}}$ (see their section 5.1), much larger than the escape velocity of the MW near the Solar circle. This is why their fig. 9 demonstrates that EMOND should have only a very small effect on the dynamics of relatively isolated galaxies, preserving the successes of standard MOND in such systems (e.g. Lelli et al. 2017). Therefore, our calculated escape velocity curve for the MW assuming standard MOND should also be very nearly correct in EMOND.

4.1 The hot gas corona of the Milky Way

4.1.1 Corona mass

It is clear that our calculated escape velocities are towards the upper end of the range allowed by observations. Thus, our analysis disfavours a hot gas corona. We have included one because *XMM–Newton* (Jansen et al. 2001) observations at a range of Galactic latitudes indicate that one is present (Nicastro et al. 2016). The mass in this corona is unclear, but their best-fitting model suggested $2 \times 10^{10}\,\mathrm{M_\odot}$ (see their table 2 model A) which we therefore use as our lowest value for M_{cor}. Similarly to this work, the best fit to their observations was obtained for the lowest mass corona model they tried out, though substantially more massive haloes are far from ruled out. A similar analysis by Miller & Bregman (2013) suggests

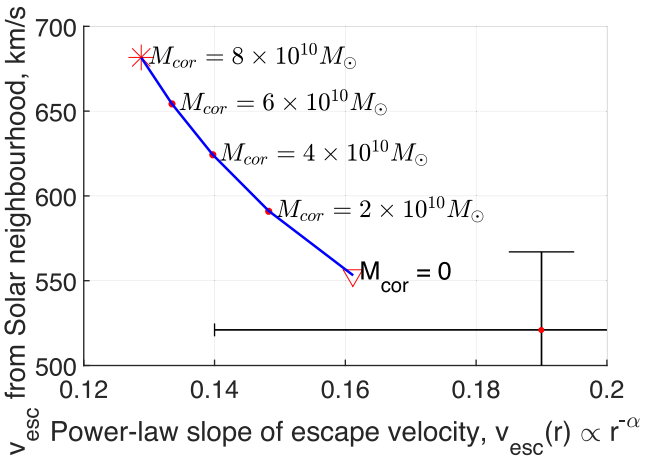


Figure 5. Effect of the MW corona mass M_{cor} on its escape velocity curve for points along its disc axis in the direction of the external field. We use the same model as in Fig. 2. The x-axis shows the value of α such that $v_{esc}(r) \propto r^{-\alpha}$ while the y-axis shows v_{esc} near the Sun. The measured values of these quantities are shown as a red dot with black error bars towards the bottom right (Williams et al. 2017).

that $M_{cor} \approx 4 \times 10^{10}\,M_\odot$, which would imply a slightly larger and slower declining v_{esc} but not to such an extent that observations rule it out (Fig. 5).

A hot gas corona would also be expected to cause ram pressure stripping effects on MW satellites containing gas. This may explain the truncation of the LMC gas disc at a much shorter distance than the extent of its stellar disc (Salem et al. 2015). Those authors used this argument to estimate that $M_{cor} = 2.7 \pm 1.4 \times 10^{10}\,M_\odot$, consistent with the other estimates.

Although our analysis is consistent with a corona of this mass, it clearly prefers an even lower mass. We therefore investigated the effect of lowering M_{cor} all the way down to 0. As expected, this would make the MW v_{esc} curve slightly more consistent with observations in terms of both its amplitude and its radial gradient (Fig. 5).

The corona of the MW would affect its satellites not only through ram pressure stripping but also by modifying their orbits, especially if the corona mass was significant compared to that of the MW stellar and gas discs. This led Thomas et al. (2017) to investigate whether a MOND model of the Sagittarius tidal stream (Newberg et al. 2002) was more consistent if a corona is included. The orbit of the progenitor would indeed be rather different with a massive corona, though the mass tried by Thomas et al. (2017) was rather high ($1.5 \times 10^{11}\,M_\odot$). This led to tidal stream radial velocities that are more consistent with observations. However, the higher MW mass also led to a much smaller apocentre for the trailing arm, making it very difficult to explain the observed distance to its apocentre (Belokurov et al. 2014). In future, a critical objective will be to see if the stars corresponding to these measurements are indeed part of the Sagittarius tidal stream, which is sometimes difficult to detect against foreground and background stars. For the time being, almost all observations of it remain consistent with a detailed MOND model of the MW that is not surrounded by a massive corona.

4.1.2 *Corona scalelength*

The scalelength of the corona is harder to pin down because spectroscopic observations are usually only sensitive to the total amount of an absorber integrated along a particular line of sight. Nonetheless, the corona can't be too centrally concentrated if it stripped the LMC gas disc (Salem et al. 2015). This suggests a scalelength comparable to the 50 kpc distance of the LMC (Pietrzyński et al. 2013). Much larger values would cause a substantial part of the corona to lie closer to M31 than to the MW (equation 14). This is unlikely given that the spatial distribution and redshifts of absorbing material in the corona strongly suggest that it surrounds the MW rather than the LG as a whole (Bregman & Lloyd-Davies 2007). This led us to explore values for r_{cor} in the range (20–60) kpc. Our results are not much affected by this parameter (Fig. 4).

4.2 The external field strength

The time-integrated effect of an EF on the LG must be that it now has a peculiar velocity v_{pec} with respect to the average matter distribution in the Universe. Assuming the direction of the EF has always remained the same, we must have that

$$\int_0^{t_f} a\, g_{ext}\, dt \;=\; v_{pec}. \qquad (23)$$

Here, t is the time elapsed since the big bang, with present value t_f. The integrating factor a is required to account for the effect of Hubble drag, which arises because objects tend to move into regions where the Hubble flow velocity is more nearly equal to the velocity of the object (e.g. Banik & Zhao 2016, equation 24). The LG currently has a peculiar velocity of $\sim$630 km s^{-1} with respect to the surface of last scattering (Kogut et al. 1993), suggesting that $g_{ext} \approx 0.015 a_0$. However, the present EF on the MW may be different to this as only its time integral is constrained and because we also need to consider other objects in the LG when estimating the EF on the MW as opposed to the EF on the LG as a whole. The most obvious such object is M31, whose sky position (Galactic co-ordinates 121.2°, −21.6°) is roughly opposite the peculiar velocity of the LG as a whole (276°, 30°). Dividing the square of the 225 km s^{-1} flatline rotation curve amplitude of M31 (Carignan et al. 2006) by its 783 kpc distance (McConnachie 2012) suggests that the resulting EF on the MW nearly cancels that due to objects outside the LG.

However, there could be non-negligible contributions arising from other objects much closer to the MW such as its satellites. The brightest MW satellite is the LMC, only $\sim$50 kpc away (Pietrzyński et al. 2013). To estimate its mass, we note that its rotation curve flatline level is $\sim$90 km s^{-1} (Kallivayalil et al. 2013), only $\sim\frac{1}{2}$ that of the MW (Kafle et al. 2012). Thus, the LMC is likely only $\sim \left(\frac{1}{2}\right)^4 = \frac{1}{16}$ as massive as the MW. The gravitational field of the LMC on the MW can be estimated by using Newton's third law, which still works in QUMOND as the theory can be derived from a least action principle (Milgrom 2010). The gravity of the MW at a distance of $r = 50$ kpc is $\sim\frac{v^2}{r}$ with $v \approx 180$ km s^{-1}, suggesting that the recoil acceleration of the MW is only $\sim 0.01 a_0$.

Of course, the LMC is not far enough away that its effect on the MW can be considered as a constant EF. Nonetheless, this gives an idea of how significant the LMC could be for our analysis, at least for v_{esc} in the Solar neighbourhood. Further away, the effect of the LMC could be larger or smaller depending on the direction. As the SDSS images the sky from the Northern hemisphere while the LMC has a declination of −70°, it should not significantly affect the v_{esc} measurements of Williams et al. (2017) that we use to constrain our models. Even so, it could lead to a larger EF on the survey volume, perhaps explaining why our models prefer slightly

higher EF strengths than suggested by equation (23). Of course, this equation does not constrain $g_{\rm ext}$ all that well because we do not know the time dependence of the EF on the LG. It could well be stronger now than its typical strength over cosmic history due to motion of objects outside the LG, e.g. if their gravitational fields on it cancelled out in the past but no longer do so as a result of recent structure formation. Thus, it is quite possible that $g_{\rm ext} \approx 0.03a_0$, as suggested by our analysis.

The high-velocity stars relevant to a $v_{\rm esc}$ measurement are expected to have a rather long orbital time and thus 'remember' conditions several Gyr ago. At that time, the smaller scalefactor of the Universe implies that large-scale structures were closer to the MW. One expects the structures themselves to be less pronounced at earlier times, but structure formation has been slowed down by the effect of dark energy and mostly happened at much earlier cosmological epochs. Thus, $g_{\rm ext}$ would likely have been larger a few Gyr ago, leading to a shallower potential well around the MW. Although this would have subsequently deepened, it may have done so too quickly for the velocity function of long-period MW stars to expand into the newly allowed region. Consequently, the observationally determined $v_{\rm esc}$ might well fall short of its true value.

4.3 The external field direction

To estimate how sensitive our results are to the direction $\widehat{\boldsymbol{g}}_{\rm ext}$ of the EF on the MW, we note that this only substantially affects the potential rather far from the MW (equation 10). At such large distances, it becomes appropriate to consider it as a point mass. This allows us to estimate how much our escape velocity calculations could be affected by $\widehat{\boldsymbol{g}}_{\rm ext}$.

We begin with a simple analytic demonstration that the effect is not likely to be large. To do this, we consider the potential difference between equally distant points located along $\widehat{\boldsymbol{g}}_{\rm ext}$ and at right angles. The distance we consider is $r_{\rm ext}$ (equation 10) because, at closer distances, the EF is not dominant. Thus, for $r \lesssim r_{\rm ext}$, we essentially have an isolated point mass, whose gravitational field is of course spherically symmetric and unaffected by $\widehat{\boldsymbol{g}}_{\rm ext}$. We use the analytic results of Banik & Zhao (2015) to obtain that

$$\Delta\Phi|_{\theta=0,\frac{\pi}{2}} = \frac{GMv_{\rm ext}}{4r_{\rm ext}} \quad \text{at} \quad r = r_{\rm ext}. \tag{24}$$

This is the amount by which the potential is expected to be deeper along $\pm\widehat{\boldsymbol{g}}_{\rm ext}$ than at right angles. However, it only applies at a distance of $r = r_{\rm ext}$ whereas we are interested in $\Delta\Phi|_{\theta=0,\frac{\pi}{2}}$ at $r = {\rm R}_\odot$. If we assume that the tangential gravity $\boldsymbol{g}_\theta$ has a similar magnitude throughout the nearly isolated region $r \lesssim r_{\rm ext}$, then this implies that

$$\Delta\Phi|_{\theta=0,\frac{\pi}{2}} = \frac{GMv_{\rm ext}{\rm R}_\odot}{4{r_{\rm ext}}^2} \quad \text{at} \quad r = {\rm R}_\odot. \tag{25}$$

For $v_{\rm esc} \sim 600\,{\rm km\,s^{-1}}$, this yields an effect of $\sim$400 m s^{-1}. In reality, $\boldsymbol{g}_\theta$ could be substantially smaller than we assumed. However, it could not be much larger because we expect $\boldsymbol{g}$ to be almost spherically symmetric near a point mass. As the radial component of $\boldsymbol{g} \lesssim r^{-1}$ for $r \lesssim r_{\rm ext}$,[3] its tangential component must increase inwards at a much lower rate, if at all. It can of course remain flat as just assumed or actually decrease inwards, e.g. $\boldsymbol{g}_\theta \propto r^1$. This yields an uncertainty of $\sim$40 times to our estimate of how much $v_{\rm esc}$ varies with $\widehat{\boldsymbol{g}}_{\rm ext}$ because ${\rm R}_\odot \approx 0.025 r_{\rm ext}$.

[3] The DML is a reasonable assumption beyond the Solar circle.

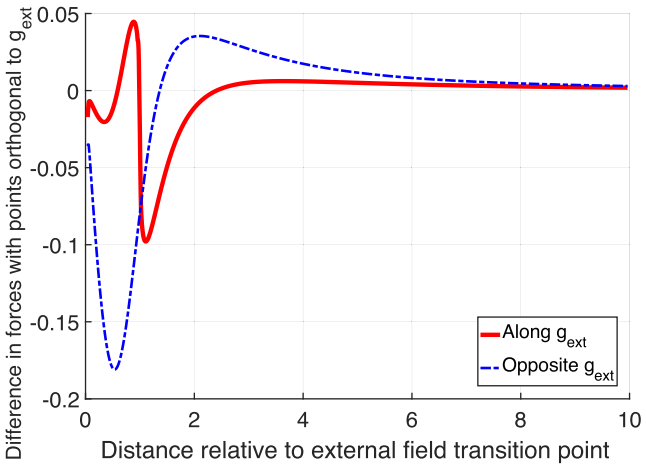


Figure 6. The difference between the radial component of the gravitational field in the directions along and orthogonal to the external field, for a point mass in the deep-MOND limit. The units are such that $G = M = a_0 = 1$. We show results for points in the direction of $\widehat{\boldsymbol{g}}_{\rm ext}$ (solid red) and $-\widehat{\boldsymbol{g}}_{\rm ext}$ (dashed blue). In both cases, the force is stronger than in the orthogonal direction at long range, as required by analytic calculations (Banik & Zhao 2015). However, this is not true at all radii. Notice how the radial forces at short range become very nearly the same in all three directions shown here, even though the gravitational field diverges as $\frac{1}{r}$. This is because the EF is unimportant at $r \ll 1$, reducing the situation to the spherically symmetric case of an isolated point mass.

Even so, it is clear that the effect of $\widehat{\boldsymbol{g}}_{\rm ext}$ does not exceed 20 km s^{-1}, a very conservative upper limit that would be correct only if $\boldsymbol{g}_\theta \propto r^{-1}$ for $r < r_{\rm ext}$. In this case, equation (24) would remain valid at the Solar circle, implying a gravitational field that deviates from spherical symmetry by $\sim\frac{1}{4}$ at $r = {\rm R}_\odot$ even though this is only a few per cent of $r_{\rm ext}$. At this position, it is highly unlikely for the gravity of a point mass to deviate this far from spherical symmetry. Nonetheless, it corresponds to a previous estimate of how much $\widehat{\boldsymbol{g}}_{\rm ext}$ might influence $v_{\rm esc}$ near the Sun (Famaey et al. 2007, section 3.2).

To obtain a better estimate, we look at the radial gravity generated by our point mass $+g_{\rm ext}$ library and integrate this to get the potential. In this case, we do not need to use a radius of $r_{\rm ext}$ but try to use ${\rm R}_\odot$ instead. Unfortunately, this is a very small fraction of $r_{\rm ext}$ and our algorithm is unable to resolve such a small scale. Thus, we use a radius of $0.04r_{\rm ext}$, which we consider acceptable because the EF is >25 times weaker than the gravity of the point mass in such regions. In reality, it is the Newtonian gravity $\boldsymbol{g}_{\rm N}$ rather than the actual gravity $\boldsymbol{g}$, which determines the phantom dark matter density that must be integrated to get $\boldsymbol{g}$ (equation 3). At these distances, $\boldsymbol{g}_{\rm N}$ from the point mass would be >625 times stronger than the EF, implying that $\boldsymbol{g}$ should be almost spherically symmetric. Moreover, the missing radial range is rather small and thus unlikely to cause a large error in the depth of the potential.

In this way, we find that $\Delta\Phi|_{\theta=0,\frac{\pi}{2}}$ corresponds to a velocity of $\sim$3 km s^{-1} which would only negligibly affect our escape velocity calculations ($\Delta v_{\rm esc} \sim 10\,{\rm m\,s^{-1}}$). To see why the effect is so small, we look at the radial component of the gravity generated by a point mass in the directions parallel and orthogonal to $\widehat{\boldsymbol{g}}_{\rm ext}$. Although the forces are stronger along $\widehat{\boldsymbol{g}}_{\rm ext}$ at large radii, they are *weaker* at small radii, leading to partial cancellation (Fig. 6). This is true both for points along $\widehat{\boldsymbol{g}}_{\rm ext}$ and in the opposite direction, albeit in slightly different ways. Therefore, we conclude that $\widehat{\boldsymbol{g}}_{\rm ext}$ should not much affect our results, making it sufficient to consider the computationally much

simpler case where $\widehat{\boldsymbol{g}}_{\rm ext}$ is aligned with the symmetry axis of the MW disc.

According to our calculations, the gravitational field due to a point mass is not symmetric between polar angles of θ and $\pi-\theta$, even if the potential very nearly is. There may be other situations where this asymmetry does have an effect. For example, a tidal stream due to a low mass MW satellite could be considered as embedded in the EF of the MW but also feeling the gravity of the progenitor. If the two are comparable, then it is likely that the tidal stream would develop asymmetrically due to the asymmetric gravitational field. For an individual object, the asymmetry could perhaps be explained in other ways such as a passing dark matter subhalo. Nonetheless, the EFE must always work in a certain way in MOND whereas a subhalo could pass by the progenitor on either side. Thus, the observation of a large number of such tidal streams could help to distinguish between the models. Unfortunately, the Sagittarius tidal stream does not seem to be much affected by the EFE in fully consistent MOND simulations of it (Thomas et al. 2017), but the theory may predict distinctive behaviour due to the EFE in other tidal streams.

5 CONCLUSIONS

In the context of MOND, we investigated the escape velocity curve of the MW using a wide range of baryonic models (Table 1). A reasonably good fit is found over the observed region (Galactocentric radii 8–50 kpc, Williams et al. 2017) using a plausible model where the MW stellar and gas discs have scalelengths fixed by observations and masses consistent with the observed MW rotation curve (Figs 3 and 4). The required external field strength is $\sim 0.03 a_0$ or slightly higher while a fairly low mass corona is preferred (Fig. 5). Indeed, our best fits are obtained if the corona does not exist at all (Fig. 5). As this is unlikely on other grounds (e.g. Nicastro et al. 2016), it is important to note that our analysis is consistent with a corona mass of $2 \times 10^{10}\,{\rm M}_\odot$, near the lower limit of the range allowed by independent observations (see their table 2 model A). Our results should not be much affected by the spatial extent of the MW corona (Section 4.1.2) or the direction of the external field on it (Section 4.3).

Using just the directly observed baryonic mass of the MW, it is possible to understand both the radial gradient of its potential and its absolute depth, as measured by its circular and escape velocity curves, respectively. It will be exciting to see whether such a model remains consistent with observations in the *GAIA* era.

ACKNOWLEDGEMENTS

IB is supported by Science and Technology Facilities Council studentship 1506672. He is grateful to Stacy McGaugh for suggesting appropriate mass models for the Milky Way. The algorithms were set up using MATLAB®.

REFERENCES

Ahn C. P. et al., 2012, ApJS, 203, 21
Banik I., 2014, preprint (arXiv:1406.4538v2)
Banik I., Zhao H., 2015, preprint (arXiv:1509.08457)
Banik I., Zhao H., 2016, MNRAS, 459, 2237
Banik I., Zhao H., 2017a, MNRAS, 467, 2180
Banik I., Zhao H., 2017b, preprint (arXiv:1701.06559)
Bekenstein J., Milgrom M., 1984, ApJ, 286, 7
Bell E. F., de Jong R. S., 2001, ApJ, 550, 212
Belokurov V. et al., 2014, MNRAS, 437, 116
Bovy J., Rix H.-W., 2013, ApJ, 779, 115
Brada R., Milgrom M., 1999, ApJ, 519, 590
Bregman J. N., Lloyd-Davies E. J., 2007, ApJ, 669, 990
Brunthaler A., Reid M. J., Falcke H., Henkel C., Menten K. M., 2007, A&A, 462, 101
Caldwell N. et al., 2017, ApJ, 839, 20
Candlish G. N., Smith R., Fellhauer M., 2015, MNRAS, 446, 1060
Carignan C., Chemin L., Huchtmeier W. K., Lockman F. J., 2006, ApJL, 641, L109
Chatzopoulos S., Fritz T. K., Gerhard O., Gillessen S., Wegg C., Genzel R., Pfuhl O., 2015, MNRAS, 447, 948
de Blok W. J. G., McGaugh S. S., 1998, ApJ, 508, 132
den Heijer M. et al., 2015, A&A, 581, A98
Famaey B., Binney J., 2005, MNRAS, 363, 603
Famaey B., McGaugh S. S., 2012, Living Rev. Relativ., 15, 10
Famaey B., Bruneton J.-P., Zhao H., 2007, MNRAS, 377, L79
Gilmore G., Reid N., 1983, MNRAS, 202, 1025
Haghi H., Bazkiaei A. E., Zonoozi A. H., Kroupa P., 2016, MNRAS, 458, 4172
Hayden M. R. et al., 2015, ApJ, 808, 132
Hodson A. O., Zhao H., 2017, A&A, 598, A127
Jansen F. et al., 2001, A & A, 365, L1
Jayaraman A., Gilmore G., Wyse R. F. G., Norris J. E., Belokurov V., 2013, MNRAS, 431, 930
Jurić M. et al., 2008, ApJ, 673, 864
Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2012, ApJ, 761, 98
Kallivayalil N., van der Marel R. P., Besla G., Anderson J., Alcock C., 2013, ApJ, 764, 161
Kogut A. et al., 1993, ApJ, 419, 1
Lehner N., Howk J. C., Wakker B. P., 2015, ApJ, 804, 79
Lelli F., McGaugh S. S., Schombert J. M., Pawlowski M. S., 2017, ApJ, 836, 152
Leonard P. J. T., Tremaine S., 1990, ApJ, 353, 486
Lüghausen F., Famaey B., Kroupa P., 2015, Can. J. Phys., 93, 232
McConnachie A. W., 2012, AJ, 144, 4
McGaugh S. S., 2008, ApJ, 683, 137
McGaugh S. S., 2011, Phys. Rev. Lett., 106, 121303
McGaugh S. S., 2016a, ApJ, 816, 42
McGaugh S. S., 2016b, ApJL, 832, L8
McGaugh S., Lelli F., Schombert J., 2016, Phys. Rev. Lett., 117, 201101
McMillan P. J., 2017, MNRAS, 465, 76
Meillon L., Crifo F., Gomez A. E., Udry S., Mayor M., 1998, in Zaritsky D., ed., ASP Conf. Ser. Vol. 136, Galactic Halos. Astron. Soc. Pac., San Francisco, p. 230
Milgrom M., 1983, ApJ, 270, 365
Milgrom M., 1986, ApJ, 302, 617
Milgrom M., 1999, Phys. Lett. A, 253, 273
Milgrom M., 2010, MNRAS, 403, 886
Miller M. J., Bregman J. N., 2013, ApJ, 770, 118
Newberg H. J. et al., 2002, ApJ, 569, 245
Nicastro F., Senatore F., Krongold Y., Mathur S., Elvis M., 2016, ApJL, 828, L12
Norris M. A. et al., 2016, ApJ, 832, 198
Olling R. P., Merrifield M. R., 2001, MNRAS, 326, 164
Ostriker J. P., Steinhardt P. J., 1995, Nature, 377, 600
Pazy E., 2013, Phys. Rev. D, 87, 084063
Peñarrubia J., Ma Y.-Z., Walker M. G., McConnachie A., 2014, MNRAS, 443, 2204
Perryman M. A. C., 1989, Nature, 340, 111
Pietrzyński G. et al., 2013, Nature, 495, 76
Piffl T. et al., 2014, A&A, 562, A91
Plummer H. C., 1911, MNRAS, 71, 460
Privon G. C., Barnes J. E., Evans A. S., Hibbard J. E., Yun M. S., Mazzarella J. M., Armus L., Surace J., 2013, ApJ, 771, 120
Quillen A. C., Garnett D. R., 2001, in Funes J. G., Corsini E. M., eds, ASP Conf. Ser. Vol. 230, Galaxy Disks and Disk Galaxies. Astron. Soc. Pac., San Francisco, p. 87

Riess A. G. et al., 1998, AJ, 116, 1009
Rubin V. C., Ford W. K., Jr, Thonnard N., 1980, ApJ, 238, 471
Salem M., Besla G., Bryan G., Putman M., van der Marel R. P., Tonnesen S., 2015, ApJ, 815, 77
Salucci P., Lapi A., Tonini C., Gentile G., Yegorova I., Klein U., 2007, MNRAS, 378, 41
Sanders R. H., 2003, MNRAS, 342, 901
Smith M. C. et al., 2007, MNRAS, 379, 755
Smolin L., 2017, preprint (arXiv:1704.00780)
Snaith O. N., Haywood M., Di Matteo P., Lehnert M. D., Combes F., Katz D., Gómez A., 2014, ApJL, 781, L31
Teyssier R., 2002, A&A, 385, 337
Thomas G. F., Famaey B., Ibata R., Lüghausen F., Kroupa P., 2017, A & A, 603, A65
Udry S. et al., 1997, in Bonnet R. M. et al., eds, ESA SP-402, Hipparcos - Venice '97. ESA, Noordwijk, p. 693
van der Marel R. P., Besla G., Cox T. J., Sohn S. T., Anderson J., 2012, ApJ, 753, 9
Verlinde E. P., 2016, preprint (arXiv:1611.02269)
Williams A. A., Belokurov V., Casey A. R., Evans N. W., 2017, MNRAS, 468, 2359
Wu X., Famaey B., Gentile G., Perets H., Zhao H., 2008, MNRAS, 386, 2199
Zhao H., Famaey B., 2012, Phys. Rev., 86, 067301
Zhao H., Famaey B., Lüghausen F., Kroupa P., 2013, A&A, 557, L3

APPENDIX A: DISCRETIZED POISSON EQUATION FOR AN EXPONENTIAL DISC

We use a spherical polar coordinate system with polar angle θ. For the axisymmetric cases we consider, we do not use the azimuthal angle. Thus, equation (5) reduces to

$$\nabla^2\Phi \;=\; \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right). \tag{A1}$$

To solve this numerically, we discretize r and θ using the integer indices i and j, respectively. We cover the range $\theta = \left(0, \frac{\pi}{2}\right)$ using 801 equally spaced points. Radially, we use a linear grid with 801 steps from $r = 0$ out to 1 and then an exponential grid (constant $\frac{r_{i+1}}{r_i}$) such that the point at $r = 1$ is halfway between the immediately adjacent radial points, minimizing sudden changes in resolution. An exponential grid at long range allows us to efficiently cover the region out to $r = 6205$ disc scalelengths, beyond which the disc potential should be very similar to that from a point mass.

The discretized version of the Laplacian operator at the cell (i, j) corresponding to $r = r_i, \theta = \theta_j$ is

$$\nabla^2\Phi = \frac{3}{{r_+}^3 - {r_-}^3}\left({r_+}^2\boldsymbol{n}_{r,+} - {r_-}^2\boldsymbol{n}_{r,-}\right) \tag{A2}$$

$$+ \frac{\sin\theta_+\boldsymbol{n}_{\theta,+} - \sin\theta_-\boldsymbol{n}_{\theta,-}}{r^2\left(\cos\theta_- - \cos\theta_+\right)}, \quad \text{where}$$

$$\boldsymbol{n}_{r,\pm} \equiv \frac{\Phi_{i\pm1} - \Phi_0}{r_{i\pm1} - r_i} \tag{A3}$$

$$\boldsymbol{n}_{\theta,\pm} \equiv \pm\frac{\Phi_{j\pm1} - \Phi_0}{r\,d\theta} \quad \text{and} \tag{A4}$$

$$r_\pm \equiv \frac{r_i + r_{i\pm1}}{2} \tag{A5}$$

$$\theta_\pm \equiv \frac{\theta_j + \theta_{j\pm1}}{2}. \tag{A6}$$

We use the convention that we only specify indices not equal to their values at the point of evaluation, where $\Phi \equiv \Phi_0$. As an example, $\Phi_{j+1} \equiv \Phi_{i,j+1}$. For the on-axis cells with $\theta = 0$, we use $\theta_- = 0$ and assume $\frac{\partial\Phi}{\partial\theta} = 0$ there due to axisymmetry. For disc cells with $\theta = \frac{\pi}{2}$, we use $\theta_+ = \frac{\pi}{2}$ and assume that

$$\frac{\partial\Phi}{\partial\theta} \;=\; r\mathrm{e}^{-\mathrm{r}} \quad \left(\theta = \frac{\pi}{2}\right). \tag{A7}$$

This imposes the surface density $\frac{\mathrm{e}^{-r}}{2\pi}$, thus creating an exponential disc with scalelength and $GM = 1$. As all the matter is in the disc, we need to solve the Laplace equation

$$\nabla^2\Phi \;=\; 0. \tag{A8}$$

The boundary condition in the disc plane imposes our desired surface density distribution. At the outermost radial shell, we expect $\Phi \approx -\frac{1}{r}$ but with a small correction due to the matter distribution having a finite extent. Taking this into account at the lowest order, we impose the boundary condition that

$$\Phi \;=\; -\frac{1}{r}\left(1 + \frac{3\left(3\sin^2\theta - 2\right)}{2r^2}\right) \quad \text{on outer shell.} \tag{A9}$$

The potential is not updated at this radial shell. To accelerate the convergence of our algorithm, we set

$$\Phi\left(r = 0\right) \;=\; -1. \tag{A10}$$

We can use equation (A2) to determine the sensitivity of $\nabla^2\Phi$ to Φ_0 and thus the value Φ_{new} such that if $\Phi_0 = \Phi_{\mathrm{new}}$, then $\nabla^2\Phi = 0$ locally. We then apply an overrelaxation method by setting

$$\Phi_0 \;\rightarrow\; W_{\mathrm{new}}\Phi_{\mathrm{new}} \;+\; \left(1 - W_{\mathrm{new}}\right)\Phi_0. \tag{A11}$$

Given n_r radial elements, we find that it is best to start with

$$W_{\mathrm{new}} \;\approx\; \frac{2}{1 + \frac{\pi}{n_r}}. \tag{A12}$$

To get our algorithm to converge, we need to gradually reduce W_{new} as it proceeds to completion, which we let it do by considering various diagnostics. We find that it is important to use $W_{\mathrm{new}} > 1.8$ because lower values lead to extremely slow progress. However, values above 2 do not work either because they lead to instability.

We apply a stringent convergence criterion to ensure that equation (A8) is satisfied everywhere subject to the boundary conditions. Equation (A2) is used to determine whether a cell has 'converged', by which we mean that $|\nabla^2\Phi| < 10^{-4}$. On a particular grid update,[4] we require convergence of all cells that are being updated (which excludes the boundary radially and the origin but not cells in the disc). The algorithm then needs to achieve this a second consecutive time, at which point we assume the solution is acceptable.

Equation (A2) implies that very tiny errors in Φ lead to large errors in $\nabla^2\Phi$ at points close to the origin. This can prevent the algorithm converging while it tries to work out the potential to, e.g. the 12th decimal place, which computers struggle to do. Thus, we increase the tolerance on $|\nabla^2\Phi|$ by a factor of $(r\sin\theta)^{-1}$ if this indeed raises the tolerance.

APPENDIX B: DISCRETIZED SOURCE FOR THE QUMOND POISSON EQUATION

The divergence of the true gravitational field $\boldsymbol{g}$ is $\nabla\cdot\left(\nu\boldsymbol{g}_{\mathrm{N}}\right)$, which we find using the same discretisation scheme as in Appendix A.

[4] That isgoing through all eight colours of our colouring scheme.

Because we need to determine ν, all components of $\boldsymbol{g}_{\mathbf{N}}$ are required at all points surrounding the one we are considering. We already have the 'divergence parts' of $\boldsymbol{g}_{\mathbf{N}}$, e.g. we know $\boldsymbol{n}_{r,\pm}$ at the points $(r_\pm, \theta)$. To get the other component of $\boldsymbol{g}_N$ at these points, we use a centred differencing scheme.

$$\frac{4\mathrm{d}\theta}{r_\pm}\left.\frac{\partial\Phi}{\partial\theta}\right|_{r_\pm} = \frac{\Phi_{i\pm1,j+1}-\Phi_{i\pm1,j-1}}{r_{i\pm1}} + \frac{\Phi_{i,j+1}-\Phi_{i,j-1}}{r_i}. \quad \text{(B1)}$$

A similar procedure would be inaccurate for $\frac{\partial\Phi}{\partial r}$ at $(r, \theta_\pm)$ due to our non-linear radial resolution scheme. Thus, we estimate $\frac{\partial\Phi}{\partial r}$ at (r, θ_j) based on fitting a parabola through Φ_{i-1}, Φ_0 and Φ_{i+1}. After applying a similar procedure to get $\frac{\partial\Phi}{\partial r}$ at $(r, \theta_{j\pm1})$, we average the results to estimate $\frac{\partial\Phi}{\partial r}$ at $(r, \theta_\pm)$.

After obtaining all components of $\boldsymbol{g}_{\mathbf{N}}$ in this way, it is simple to multiply each potential derivative appearing in equation (A2) by the value of ν at the corresponding position, thus yielding the divergence of the true gravitational field.

This paper has been typeset from a TEX/LATEX file prepared by the author.